

Checking the Math: Government Secrecy and DNA Databases

SARAH M. RUBY*

Abstract: Popular television shows such as CSI portray DNA evidence as the equivalent of a positive identification. In reality, DNA matches are as significant as statistics tell us they are. DNA match statistics are calculations that describe the likelihood of a coincidental match between a person suspected of crime and the DNA sample found at a crime scene. These numbers are difficult for laypeople to understand, and the methodology behind them is the subject of scientific debate. One way to study the reliability of DNA evidence is to examine criminal offender DNA databases. These stores of genetic data are much larger than the databases scientists have used to develop DNA statistics. But offender DNA databases are secret, government agencies use them to find suspects, and only their own technicians have access to them.

Arguments in support of database secrecy do not hold up. Government agencies are accustomed to handling and disseminating all kinds of sensitive information about individuals. This article examines three contexts in which the government discloses sensitive information: public records releases, court trials, and medical research. The

* J.D., University of California, Berkeley School of Law (Boalt Hall), expected 2010; B.A., Stanford University, 2001. I am deeply grateful to Professor Erin Murphy, whose insights and encouragement made this article possible, and to Deputy Public Defender Bicka Barlow for inspiring its writing. Thanks also to Special Litigation and Forensics Fellow Katherine Philpott, Deputy Attorney General Michael Chamberlain, Instructor Ron C. Michaelis, Professor Laurence D. Mueller, Professor Franklin E. Zimring, Roxana Killian, Joanna Lydgate, and Jonas Lerman for their thoughtful comments. I am indebted to Alan Lagod for taking an interest in this research, and to staff members of *I/S: A Journal of Law and Policy for the Information Society* for their careful editing. And to Mark Allen, thank you most of all.

most appropriate way of handling offenders' DNA profiles, this article argues, is to follow practices prevalent in the medical-research context. Government agencies routinely collect DNA samples from newborn babies and share them with medical researchers. If this system adequately protects children's genetic information, it could also safeguard criminal offenders' DNA profiles. If government agencies do not allow scientists to test the assumptions underlying DNA evidence, they risk violating the constitutional rights of criminal defendants and undermining the credibility of DNA typing as a forensic method.

INTRODUCTION

Studies show that DNA evidence is particularly persuasive to juries.¹ It has a "special aura of certainty and mystic infallibility,"² and "may be so persuasive that its mere introduction in a criminal case is sufficient to seriously impede defense challenges."³

DNA helped convict John Puckett of murder in 2008.⁴ The San Francisco Police Department's crime lab had identified him by searching California's offender DNA database, matching his DNA to a trace of sperm left at a violent murder scene in 1972.⁵ Puckett, an aging sex offender, argued that he did not kill anyone, that the DNA match reflected a mistake or coincidence.⁶ At his trial, a prosecution witness told the jury that the chance of Puckett's DNA coincidentally matching the true killer's sperm was one in 1.1 million.⁷ Puckett's

¹ Joel D. Lieberman et al., *Gold Versus Platinum: Do Jurors Recognize the Superiority of DNA Evidence Compared to Other Types of Forensic Evidence?* 14 PSYCHOL. PUB. POL'Y & L. 27 (summarizing three studies examining the influence of DNA evidence on juries).

² *Id.* at 52.

³ *Id.* at 58.

⁴ Jason Felch & Maura Dolan, *DNA: Genes as Evidence: The Odds of Justice Can Be Long*, L.A. TIMES, May 4, 2008, at A1; Chris Smith, *DNA's Identity Crisis*, S.F. MAG., Sept. 2008, at 74.

⁵ Felch & Dolan, *supra* note 4; Smith, *supra* note 4.

⁶ Felch & Dolan, *supra* note 4.

⁷ Trial Transcript at 2165, *People v. Puckett*, No. SCN 201396 (Cal. Super. Ct. Jan. 22, 2008) (Bonnie Cheng, a criminalist with the San Francisco Police Department Crime Lab, testified that "the probability that a random unrelated individual by chance would possess

defense expert, using a formula endorsed by the FBI and the National Research Council, put the chance of a coincidental match at one in three.⁸ The judge did not allow the jury to hear this number.⁹ Yet if it is correct, San Francisco's crime lab would have had a better than thirty percent chance of finding a match in the database regardless of whether the killer's DNA profile was actually stored there.¹⁰

the same DNA profile as that detected in the sperm fraction of the oral smear slide is approximately one in 1.1 million for U.S. Caucasians.”). Puckett's case is a good example of how slippery these numbers can be for lawyers and jurors. Even though the prosecution's witness characterized her rarity calculation accurately—as the chance of a coincidental match between Puckett and the actual perpetrator—the prosecutor misapplied it during closing argument. He contended that the denominator of the random match probability (1.1 million) was equal to the number of people who must be tested in order to find another matching profile. *Id.* at 3363. Thus of the approximately nine million white men in California at the time of the murder, he argued that only “eight to nine . . . men . . . look like that crime scene profile and [Puckett] happened to be one of them.” *Id.* at 3364. This is a common “prosecutor's fallacy,” and it “overstates the strength of the evidence because it overestimates the number of people who must be tested before another match is likely to be seen.” RON C. MICHAELIS ET AL., *A LITIGATOR'S GUIDE TO DNA* 98 (2008).

⁸ Trial Transcript at 2833–34, *People v. Puckett*, No. SCN 201396 (Cal. Super. Ct. Feb. 4, 2008) (Laurence D. Mueller, professor at UC Irvine, said that because Puckett was identified through a database search, “the [correct] statistic would be about one in three.”) A couple of issues are worth noting here. First, the opposing experts' numbers are both statistically defensible—they simply use different assumptions. MICHAELIS ET AL., *supra* note 7, at 91 (explaining random match probability generally). How to figure out whether their assumptions are correct is the heart of this article. Second, the defense expert made his calculation by following the National Research Council's and FBI's recommendation for modifying DNA statistics where a suspect is located through a database search. NAT'L RESEARCH COUNCIL, *THE EVALUATION OF FORENSIC DNA EVIDENCE* 40, 134 (1996); FBI DNA Advisory Board, *Statistical and Population Genetics Issues Affecting the Evaluation of the Frequency of Occurrence of DNA Profiles Calculated from Pertinent Population Database(s)*, FORENSIC SCI. COMM. Feb. 23, 2000, *available at* <http://www.fbi.gov/hq/lab/fsc/backissu/july2000/dnastat.htm>. This modification is known as the NRC II formula, and in most cases, it is calculated by multiplying the random match probability (the prosecutor's statistic) by the number of profiles in the database. *Id.* Using NRC II, Mueller multiplied the prosecution's calculation—one in 1.1 million—by 300,000, which was the number of profiles California stored in its offender database at the time. Smith, *supra* note 4, at 76.

⁹ Trial Transcript at 2852, *People v. Puckett*, No. SCN 201396 (Cal. Super. Ct. Feb. 4, 2008).

¹⁰ When San Francisco's crime lab hit on Puckett's profile, the California database contained some 300,000 DNA profiles. Smith, *supra* note 4, at 76. Today, California's offender database contains more than 1 million profiles. FBI, CODIS-NDIS Statistics, <http://www.fbi.gov/hq/lab/codis/stats.htm> (last visited Apr. 16, 2010) [hereinafter CODIS Statistics]. To calculate the chance of finding a match in today's database, the NRC II formula must be adjusted because multiplying the random match probability (one in 1.1 million) by the size of the DNA database (1 million) approaches one, or 100 percent. E-mail from Laurence D. Mueller, Professor of Ecology and Evolutionary Biology, Univ. of Cal.,

Statistics are what make DNA evidence probative in court.¹¹ DNA statistics describe the likelihood that a match between crime-scene evidence and a potential perpetrator is coincidental.¹² While these statistics do not reflect the incidence of matching profiles in the population, in most cases the lower the chance of a coincidental match, the rarer the DNA profile.¹³ Just as a Social Security number would be a poor identifier if many people shared the same one, a DNA profile must be uncommon to be probative of guilt. Current DNA statistics suggest that DNA profiles are rare, but scientists developed those statistics from small population-genetics databases that generally contain about 1000 people.¹⁴ Scientists debate the methodology behind DNA statistics, with many arguing that empirical study of large numbers of DNA profiles is the best way to establish whether the assumptions they rely upon are correct.¹⁵

Puckett was 72 years old and living with his wife in a trailer park eighty miles from San Francisco when police arrested him for the decades-old murder of Diana Sylvester, a 22-year-old nurse.¹⁶

Irvine, to author (July 27, 2009 19:27 PST) (on file with author). The adjusted chance of finding a coincidental match in today's database is roughly one in 1.6, or 65 percent. *Id.*

¹¹ See, e.g., *United States v. Davis*, 40 F.3d 1069, 1075 (10th Cir. 1994) (noting that "statistical probabilities are basic to DNA analysis"); *United States v. Coronado-Cervantes*, 912 F. Supp. 497, 500 (D.N.M. 1996) (explaining that once a DNA match is made, "[s]tatistics must then be generated to give significance to the match"). See also *Young v. State*, 879 A.2d 44, 44–45 (2005) (holding that evidence of a DNA "match" is admissible without statistics, but resting this holding on the fact that the DNA method analyzed "sufficient locations to arrive at an infinitesimal random match probability").

¹² MICHAELIS ET AL., *supra* note 7, at 91

¹³ E-mail from Ron Michaelis, Instructor, Dep't of Biology, W. Carolina Univ., to author (May 3, 2010 11:37 PST) (on file with author).

¹⁴ See, e.g., Letter from Dan E. Krane, Professor of Biological Sciences, Wright State Univ., et al., to Christian Hassell, Dir., FBI Laboratory (Oct. 15, 2009) (on file with author).

¹⁵ See, e.g., D. E. Krane et al., *Time for DNA Disclosure*, 326 SCIENCE 1361 (2009); Decl. of Laurence D. Mueller, *United States v. Berger*, No. 2004 FEL 003420, at 5 (D.C. Super. Ct. Oct. 22, 2006) [hereinafter Mueller Decl. II]; Decl. of Laurence D. Mueller, *United States v. Berger*, No. 2004 FEL 003420, at 1 (D.C. Super. Ct. Oct. 6, 2006) [hereinafter Mueller Decl. I]; Letter from Keith Devlin, *People v. Ford*, No. SCN 2151192 (Cal. Super. Ct. Oct. 3, 2008); Letter from Donald Kennedy, *People v. Ford*, No. SCN 2151192 (Cal. Super. Ct. Oct. 3, 2008); Decl. of Montgomery Slatkin, *People v. Ford*, No. SCN 2151192 (Cal. Super. Ct. Oct. 3, 2008); Letter from Dan E. Krane et al., *supra* note 14.

¹⁶ Felch & Dolan, *supra* note 4; Smith, *supra* note 4, at 72, 74.

Sylvester died a few days before Christmas in her San Francisco apartment, where she had been sexually assaulted, strangled, and stabbed in the heart.¹⁷ Her landlord had seen the killer—she heard noises and entered the apartment’s open front door.¹⁸ A man hustled her out, growling, “Go away, we’re making love.”¹⁹

The killer’s identity remained a question for more than thirty years. It might have stayed that way if not for forensic DNA evidence,²⁰ which was relatively weak in Puckett’s case. The crime-scene DNA sample came from the inside of Sylvester’s mouth, which meant it contained a mixture of the victim’s and her killer’s DNA.²¹ Furthermore, the crime-scene sample had decayed such that fewer than half of the typical thirteen DNA markers could be identified.²² The non-DNA evidence against Puckett was circumstantial: he had lived in San Francisco at the time of the murder, committed several sex offenses there five years later, and been referred for employment

¹⁷ Felch & Dolan, *supra* note 4; Smith, *supra* note 4, at 72.

¹⁸ Smith, *supra* note 4, at 72.

¹⁹ *Id.*

²⁰ See Press Release, S.F. Police Dep’t, Arrest Through DNA Analysis in 33-Year-Old Homicide Case (Apr. 21, 2006), available at <http://sf-police.org/index.aspx?page=2113>.

²¹ Smith, *supra* note 4, at 80; Trial Transcript at 2125, *People v. Puckett*, No. SCN 201396 (Cal. Super. Ct. Jan. 22, 2008). The significance of DNA mixtures is controversial. Bonnie Cheng, a criminalist at the San Francisco Police Department Crime Lab, testified in Puckett’s case that a simple process separates sperm cells from non-sperm cells before profiling each. *Id.* at 2116. However, she also testified that it’s fairly common for the separation to be incomplete, leaving DNA technicians, and ultimately juries, to guess which profiled DNA markers belong to the killer and which belong to the victim. *Id.* at 2121. No established protocol exists for how to interpret DNA mixtures, and Peter Gill, a British researcher who administers Britain’s national DNA database, told a forensic science conference in 2005 that “[i]f you show 10 colleagues a mixture, you will probably end up with 10 different answers.” Smith, *supra* note 4, at 80. In Puckett’s case, some of the DNA attributed to the killer could have been Sylvester’s, leading police to look for a DNA profile in the database that was not necessarily consistent with the killer’s DNA profile. See Trial Transcript at 2212, *People v. Puckett*, No. SCN 201396 (Cal. Super. Ct. Jan. 22, 2008).

²² Felch & Dolan, *supra* note 4; Smith, *supra* note 4. In another DNA case in Illinois, the state database administrator testified in a deposition that “[i]f it doesn’t match across the thirteen [markers], then it’s not a true match.” *People v. Wright*, Nos. 1-07-3106, 1-07-3464, 2010 WL 1194903, at *8 (Ill. App. Ct. Mar. 26, 2010) (referring to a deposition taken in connection with *People v. Luna*, No. 02 CR 15430 (Ill. Cir. Ct. 2006)). Puckett matched the crime-scene sample at five-and-a-half markers. Felch & Dolan, *supra* note 4; Smith, *supra* note 4.

to the University of California, San Francisco, where the victim worked as a nurse.²³ After the trial, jurors said it was the DNA—and particularly the prosecution's statistic—that convinced them Puckett was guilty.²⁴

These seemingly irreconcilable statistics offered by opposing experts in the Puckett trial highlight one of the reasons scientists urge empirical study of DNA profiles—they question the role DNA databases play in increasing the likelihood of coincidental matches.²⁵ In Puckett's case, the prosecution told the jury that the chance of a coincidental match was one in 1.1 million, and the defense countered with one in three.²⁶ It did not come up at trial, but a third school of scientists would say the prosecution's statistic was too conservative—in their view, the odds would be longer that Puckett coincidentally matched Sylvester's killer because running a database search rules out thousands of potential suspects.²⁷ Interestingly, these conflicting views—one insisting that a database search makes a false match more likely, and the other contending it does the exact opposite—"rest solidly on valid mathematical principles."²⁸ Both are correct, mathematically speaking.²⁹

²³ Trial Transcript at 3365–69, *People v. Puckett*, No. SCN 201396 (Cal. Super. Ct. Feb. 13, 2008); Felch & Dolan, *supra* note 4.

²⁴ Felch & Dolan, *supra* note 4. "I don't think we'd be here if it wasn't for the DNA," a juror told the *Los Angeles Times*. *Id.* When asked if the defense's one in three statistic would have affected his decision, he said, "Of course it would have changed things." *Id.*

²⁵ See, e.g., Mueller Decl. I, *supra* note 15.

²⁶ For an explanation of how each side reached its statistic, see *supra* note 8.

²⁷ See Peter Donnelly & Richard D. Friedman, *DNA Database Searches and the Legal Consumption of Scientific Evidence*, 97 MICH. L. REV. 931, 931, 945 (1998–1999) (arguing that finding a DNA match within a database is slightly more probative than a random non-database match because it eliminates as potential suspects the many people within the database whose DNA profiles do not match the crime-scene evidence).

²⁸ MICHAELIS ET AL., *supra* note 7, at 127; see also FBI DNA Advisory Board, *supra* note 8 ("Both camps appear to present rigorous arguments to support their positions. Indeed the proper treatment superficially appears to rest in the details of arcane mathematics.")

²⁹ That the scientific community could entertain diametrically opposing views "seems surprising [to laypeople], but in fact in the history of mathematics, and in particular the history of probability theory, there have been several cases where it took some time before a consensus emerged as to what calculation was appropriate in a particular situation." Letter from Keith Devlin, *supra* note 15.

While conflicting truths and questionable assumptions are tolerable in the abstract, they become troubling in a criminal justice system that operates on reasonable doubt. The statistical and genetic questions raised by Puckett's case and others cast doubt upon the probative value of all DNA matches, whether or not they involve databases.³⁰ A way to find out the true significance of a DNA match, some scientists say, is to study the contents of offender databases, which contain far more profiles than any database previously studied.³¹ However, these databases are largely secret. Government agencies search them to find suspects and regard all other uses to be illegal. Using these and other arguments, government lawyers often persuade judges to deny outsiders access to the contents of DNA databases.³² This article evaluates the arguments for and against

³⁰ The effect of a database search on the likelihood of finding a DNA match is just one question raised by the use of forensic DNA evidence. Another concern is that DNA statistics might incorrectly assume that all DNA markers are independent, discounting the possibility that certain combinations of markers might be more common than others. Laurence D. Mueller, *Can Simple Population Genetic Models Reconcile Partial Match Frequencies Observed in Large Forensic Databases?*, 87 J. GENETICS 101, 107 (2008). Others can be found *infra* notes 91–94 and surrounding text.

³¹ *E.g.*, Letter from Dan E. Krane et al., *supra* note 14 (sixteen academics asking for access to the national DNA database due to concern about the accuracy of DNA rarity statistics).

³² See *United States v. Blackmon*, No. 2008 CF1 21355 (D.C. Super. Ct. Feb. 23, 2009) (Alprin, J.); *United States v. Berger*, No. F-3420-04 (D.C. Super. Ct. July 14, 2006) (Satterfield, J.); *United States v. Jacobs*, No. F-320-00 (D.C. Super. Ct. Feb. 7, 2006) (Reid-Winston, J.); *United States v. Rue*, No. F-3817-05 (D.C. Super. Ct. Dec. 29, 2005) (Christian, J.); *People v. Ford*, No. SCN 2151192 (Cal. Super. Ct. July 29, 2008); Notice of Objection and Motion to Quash Subpoena Duces Tecum, *People v. Ford*, No. SCN 2151192, at 2, fn.1 (Cal. Super. Ct. Aug. 5, 2008) [hereinafter *Ford Motion to Quash*] (citing *Davis v. Super. Ct.*, No. A116603 (Cal. Ct. App. Mar. 8, 2007); *People v. Brown*, No. CC643081 (Cal. Super. Ct. July 11, 2008); *People v. Davis*, No. SCN 190226 (Cal. Super. Ct. Nov. 21, 2006) (Morgan, J.); *People v. Perryman*, No. SCN 198604 (Cal. Super. Ct. Nov. 16, 2006); *People v. Smith*, No. BA 255390 (Cal. Super. Ct. Sept. 26, 2006); *People v. Ardoin*, No. 196690-3 (Cal. Super. Ct. Sept. 20, 2006); *People v. Hill*, No. KA 064034 (Cal. Super. Ct. Feb. 24, 2005)). *But see* *People v. Wright*, Nos. 1-07-3106, 1-07-3464, 2010 WL 1194903 (Ill. App. Ct. Mar. 26, 2010) (unanimous appellate panel ordered new trial for defendant denied search of DNA profile matches within Illinois's database); *People v. Luna*, No. 02 CR 15430 (Ill. Cir. Ct. 2006) (judge ordered search for DNA profile matches within Illinois's database); *State v. Davis*, No. CT 05-0024X (Md. Cir. Ct. Aug. 4, 2006) (Platt, J.) (judge ordered search for DNA profile matches within Maryland's database); E-mails from Steve Jacobson, Attorney, to author (Mar. 12, 2010 11:19 & 14:48 PST) (on file with author) (referring to a 1999 order in *People v. Shreck*, No. 98 CR 2475, in Boulder, Colorado, in which Judge Daniel Hale ordered Colorado to make the contents of its offender database available to Steve Jacobson, Shreck's attorney; at the time, the database contained 5,629 profiles).

database secrecy. It explores ways of making offender DNA databases available without compromising the sensitive data they contain and concludes that independent scientific researchers should be allowed to study them.

This proposal is modest.³³ A more radical one would ask the government to hand over offenders' DNA samples rather than profiles.³⁴ DNA samples are the biological material itself, while DNA profiles are twenty-six-number³⁵ sequences that derive from DNA samples and correspond to a tiny portion of the full human genome.³⁶ A study of DNA samples would let researchers check the work of government laboratories, a task past experience suggests is

³³ It is also gaining traction among criminology scholars. William C. Thompson, professor and chair of the Department of Criminology, Law & Society at the University of California, Irvine, writes that "there is no persuasive justification for the government's insistence on maintaining the secrecy of database profiles, so long as the identity of the contributors is not disclosed." William C. Thompson, *The Potential for Error in Forensic DNA Testing*, 21 GENEWATCH 5, 8 (2008). Likewise, Professor David Kaye writes that "the government should make an anonymized version of [the national database] available to all researchers. The ultimate outcome . . . will be greater confidence in the method now used . . . , or it will be some revised, but more defensible form of these estimates." David H. Kaye, *DNA Database Woes: What is the FBI Afraid Of?* 31–32 (unpublished manuscript, on file with author).

³⁴ DNA sample research is underway in the private sector, as companies capitalize upon people's desire to unlock the secrets of their DNA. *Spitting Parties: The Fashionable Approach to Genetic Risk?*, LAB TIMES ONLINE, Oct. 1, 2008, http://www.lab-times.org/editorial/e_078.html. Companies such as 23andMe encourage customers to host "Spit Parties," in which guests submit a saliva sample for analysis. *Id.* Rupert Murdoch and Harvey Weinstein are "[k]nown spitters." *Id.* They and other "[c]ustomers cannot opt out of having their information anonymously shared, but they can refuse to participate in surveys focusing on specific traits." *Id.* People that seek out these tests tend to be wealthy and might not worry about genetic discrimination in health care and other areas. However, the fact that people surrender their entire genome to private companies suggests that fears about the misuse of forensic DNA profiles are overblown.

³⁵ Law enforcement agencies once used fewer numbers, and some use more, but the FBI has settled on twenty-six as standard. BRUCE BUDOWLE ET AL., CODIS AND PCR-BASED SHORT TANDEM REPEAT LOCI: LAW ENFORCEMENT TOOLS 76 (1998), available at <http://www.promega.com/geneticidproc/eusymp2proc/17.pdf>. Twenty-six numbers correspond with thirteen DNA markers, each of which contain two alleles.

³⁶ See MICHAELIS ET AL., *supra* note 7, at 11–12, 16. Forensic testing focuses on portions of the genome known as "benign polymorphisms," which seem to have no impact on the function of a gene's proteins and therefore no impact on human health. *Id.* at 11. Portions of the genome known to affect genes' proteins—"functional polymorphisms"—"play critical roles in our development and function" and have been linked to diseases such as cancer, Alzheimer's disease, diabetes, asthma, and cardiovascular disease. *Id.* at 11–12.

warranted.³⁷ For example, an FBI DNA lab analyst, Jacqueline Blake, pled guilty to falsifying her laboratory's quality assurance reports.³⁸ An audit of a Massachusetts crime lab uncovered faulty DNA test results and instances in which someone had mistakenly entered the same genetic profile for two individuals.³⁹ And the crime lab that identified Puckett was faulted for substandard cleanliness and inadequate chain-of-custody records.⁴⁰ These and other laboratory failures reveal the vulnerabilities inherent in our system of collecting, identifying, and analyzing DNA samples.⁴¹ The study of offenders' DNA samples would be a check against systemic problems and "horrific tales of false-positive DNA matches," which touch some of the nation's premier DNA laboratories.⁴² Furthermore, the release of DNA samples would let researchers develop a clean database of DNA profiles for population-genetics research—a purpose for which government scientists say existing profile databases are ill-suited.⁴³

Yet DNA samples are revealing. They contain entire genomes, and they might reveal the genetic predispositions of a given offender population. Profiles, on the other hand, are strings of digits. They represent a portion of the genome known as "junk DNA"—so-named because they are not believed to code for characteristics. In the digital

³⁷ See Erin Murphy, *The New Forensics: Criminal Justice, False Certainty, and the Second Generation of Scientific Evidence*, 95 CAL. L. REV. 721, 773 (2007). False positives can result from contamination and poor quality control, such as improper use of gloves or "squeezing a pipette into the wrong tube." *Id.* at 754. DNA is also potentially vulnerable to misuse "by rogue police officers doggedly committed to obtaining convictions." Troy Duster, *Explaining Differential Trust of DNA Forensic Technology: Grounded Assessment or Inexplicable Paranoia?* 34 J. L. MED. & ETHICS 293, 294 (2006). "Among the first 200 people exonerated by post-conviction DNA testing were two men . . . who were convicted in the first place due partly to DNA testing errors." Thompson, *supra* note 33, at 7.

³⁸ Murphy, *supra* note 37.

³⁹ *Id.*

⁴⁰ Am. Soc. Of Crime Laboratory Dirs., Inspection Report 5, 20 (Nov. 2009). A San Francisco crime lab employee was also accused of stealing drug evidence. Jaxon Van Derbeken, *Crime Lab Fallout: Drug Defendants Go Free*, S.F. CHRON., May 11, 2010, available at <http://www.sfgate.com/cgi-bin/article.cgi?f=/c/a/2010/03/11/MNG81CDOTU.DTL>.

⁴¹ MICHAELIS ET AL., *supra* note 7, at 92.

⁴² Murphy, *supra* note 37, at 754–55.

⁴³ Bruce Budowle et al., *Partial Matches in Heterogeneous Offender Databases Do Not Call into Question the Validity of Random Match Probability Calculations*, 123 INT. J. LEGAL MED. 59, 63 (2009).

era, a DNA profile is less revealing than a Social Security number.⁴⁴ If scientists discover that forensic DNA profiles code for traits, the most we will learn is that offenders do or do not have that trait. In a society that worries about supermarket smart cards and Google's use of search terms,⁴⁵ concern about the possible misuse of an immutable string of numbers is understandable. But unlike Google and the supermarket, which put personal data to economic use, government agencies use DNA profiles to identify suspects in criminal cases.

This article calls for researcher access to offenders' profiles rather than samples.⁴⁶ The low-hanging fruit in the forensic DNA debate is whether rarity statistics presented in court are accurate. Statisticians, population geneticists, and other academics say they can answer this question using existing stores of offenders' DNA profiles.⁴⁷ Respect for defendants' due process and other constitutional rights require that they be allowed to try.⁴⁸

Though questions about the rarity of DNA profiles arise most often in the context of a criminal trial, the adversarial criminal justice system is not the ideal forum for answering them.⁴⁹ If U.S. government agencies continue to resist any outside testing of their

⁴⁴ "Identity thieves can use your [Social Security] number and your good credit to apply for more credit in your name." Social Security Admin., Identity Theft and Your Social Security Number, <http://www.ssa.gov/pubs/10064.html> (last visited Apr. 16, 2010). DNA profiles are currently useless to identify thieves.

⁴⁵ See, e.g., Consumers Against Supermarket Privacy Invasion and Numbering, Food Industry News, <http://www.nocards.org/news/archive2.shtml> (last visited Apr. 15, 2010); Alex Pham & Michelle Quinn, *Critics to Google: Privacy Please*, L.A. TIMES, Mar. 15, 2007, at C1.

⁴⁶ Research on offenders' DNA samples might someday prove necessary. However, fears that the study of offender DNA profiles will lead to this research is not a reason to avoid the more limited step.

⁴⁷ See, e.g., D. E. Krane et al., *supra* note 15; Letter from Dan E. Krane et al., *supra* note 14 (asking, as sixteen academics concerned about the accuracy of DNA rarity statistics, for access to the national DNA database).

⁴⁸ The nature of potential constitutional violations stemming from DNA database secrecy is beyond the scope of this article. However, in a recent subpoena of the California offender database, a defense lawyer invoked her client's rights to confrontation, cross-examination, and due process. Opposition to Attorney General's Motion to Quash Subpoena Duces Tecum, *People v. Ford*, No. SCN 2151192 (Cal. Super. Ct. Aug. 15, 2008); See also Murphy, *supra* note 37, at 791 (noting that defendants' counsel and due process rights "clearly contemplate" at least some access to offender databases).

⁴⁹ See *infra*, notes 359–366 and accompanying text.

databases,⁵⁰ however, the courtroom remains the best venue in which to address the constitutionality of total secrecy. Defense lawyers will continue to seek access to databases, or ask to run special searches. These requests are seldom successful,⁵¹ but they show no sign of flagging even as the media begins to call for outside testing of offender databases.⁵² Government agencies have the power to put a stop to this type of litigation, or at least buttress its arguments against it. Existing laws allow them to release an anonymous, encrypted version of the database to independent statistical researchers.⁵³

If the government chooses secrecy, forensic DNA technology could lose its place as the gold standard of forensic methods. In early 2009, the National Academy of Sciences (NAS) slammed as unscientific every forensic technique other than DNA evidence.⁵⁴ With the exception of DNA analysis, the NAS found, “no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source.”⁵⁵ Even so, the NAS noted that “DNA evidence is [not] always unassailable in the courtroom.”⁵⁶ For example, searching large DNA databases for suspects “might yield false positives with some regularity.”⁵⁷ Pretending these problems do

⁵⁰ For example, the FBI recently denied a request from sixteen statisticians, geneticists, and other academics to study the national offender database. Letter from D. Christian Hassell, Dir. FBI Laboratory, to Dan E. Krane, Professor of Biological Sciences, Wright State Univ. (Dec. 22, 2009) (on file with author). However, it recognized the need for “revalidating previous assumptions and continuing to improve the foundations upon which the database is based...” *Id.* at 2. To that end, the FBI is “exploring ways to address some of the topics” the scientists raised. *Id.*

⁵¹ See *supra* note 32.

⁵² See, e.g., Editorial, *DNA Evidence: What Are the Real Chances of Mistakes?*, LAS VEGAS REV. J., July 29, 2008, at B10; Editorial, *Should We Trust DNA?*, S.F. CHRON., July 28, 2008, at B4.

⁵³ See *infra* Part III.

⁵⁴ NAT’L RES. COUNCIL COMM. ON IDENTIFYING THE NEEDS OF THE FORENSIC SCIS. CMTY., *STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 7* (2009), available at http://www.nap.edu/catalog.php?record_id=12589 [hereinafter NAS Report].

⁵⁵ *Id.*

⁵⁶ *Id.* at 100.

⁵⁷ *Id.* at 100 n.61 (summarizing the work of B.S. Weir, *The Rarity of DNA Profiles*, 1 ANNALS APPLIED STAT. 358 (2007)).

not exist will not make them go away. By hiding DNA from rigorous scientific scrutiny, the government risks discrediting the one forensic technique rooted in science.⁵⁸ Donald Kennedy, a biologist at Stanford University and co-chair of the committee that produced the NAS Report, advocates for scientific study of offender databases.⁵⁹ “Science depends on the free flow of information and data,” he wrote in a letter supporting a defense attorney’s attempt to subpoena California’s offender database.⁶⁰ “To prevent access to, and analysis of, these data would be inappropriate and antithetical to scientific progress.”⁶¹ Science requires inquiry even when the outcomes are uncertain or likely to confirm conventional wisdom.⁶²

People are going to prison based on a technology that requires further study, and government agencies control the only data set large enough to check the math involved. While this article discusses this issue as it plays out in the federal and state governments, it primarily focuses on California, home to the largest state offender database in the nation.⁶³ Part I describes the ways in which law enforcement uses DNA databases, and the scientific debate regarding prevailing DNA statistics. Part II evaluates arguments used by government agencies to block outsider access to DNA databases. Part III examines DNA database statutes and determines that they allow the disclosure of DNA profiles for scientific research. Part IV examines three situations in which government agencies disclose sensitive information about individuals, and calls for the release of DNA profiles to qualified researchers using controls already in place for genetic health research.

⁵⁸ *Id.* at 8 (noting that “there is a notable dearth of peer-reviewed, published studies establishing the scientific bases and validity of many forensic methods”).

⁵⁹ Letter from Donald Kennedy, *supra* note 15 (pointing out that “[t]here is considerable scientific interest in this problem—not only because it might help clarify an unexpected result, but because the knowledge to be gained may assist in the appropriate use of such matching techniques in judicial proceedings.”).

⁶⁰ *Id.*

⁶¹ *Id.*

⁶² *Id.*

⁶³ CODIS Statistics, *supra* note 10.

I. THE PROBLEM OF DNA DATABASE SECRECY

California's DNA database dates to 1983, when the legislature passed a law requiring convicted sex offenders to provide blood and saliva specimens upon release from prison.⁶⁴ Other states also began creating DNA databases around this time,⁶⁵ and in 1994 Congress passed the DNA Identification Act, which established the FBI's Combined DNA Index System (CODIS) linking local, state, and federal DNA databases.⁶⁶ The DNA profiles in these databases fall into two general categories: (1) "forensic" profiles derived from DNA found at crime scenes, and (2) "offender" or "known" profiles, which correspond to persons who voluntarily donated DNA samples or were forced to do so by statute.⁶⁷

CODIS, which went live in October 1998,⁶⁸ linked state and federal databases that would grow to include more than 7.9 million offender DNA profiles in eleven years.⁶⁹ California is by far the biggest

⁶⁴ CAL. PENAL CODE § 290.2, *repealed by* DNA and Forensic Identification Database and Data Bank Act of 1998, CAL. PENAL CODE § 295 (West 2010).

⁶⁵ In 1985, England was the first country to use forensic DNA for law-enforcement purposes. H.R. Rep. No. 103-45 (1993). Florida was the first state to follow suit, in 1987. *Id.*

⁶⁶ DNA Identification Act of 1994, Pub. L. No. 103-322, 108 Stat. 2065; Peter Finn, *Revolution Underway in Use of DNA Profiles*, WASH. POST, Nov. 16, 1997, at B4. The crime-fighting potential of DNA was incredibly promising, but some were wary of a centralized genetic repository. *See, e.g.*, Robert S. Boyd, *DNA on File for Millions in America—States Saving Data on Babies, Criminals; Privacy Fears Raised*, SEATTLE TIMES, Nov. 1, 1994, at A12. James Watson, who shared the Nobel Prize for helping discover DNA, lobbied against genetic data collection, testifying at a 1991 congressional hearing that "[t]he idea that there will be a huge databank of genetic information on millions of people is repulsive." *Id.*

⁶⁷ *See* Murphy, *supra* note 37, at 738. Professor Murphy notes that scholars and courts continue to grapple with the scope and constitutionality of the many offender-collection statutes on the books. Murphy, *supra* note 37, at 738 n.73. Courts have generally upheld mandatory DNA sampling of convicted offenders. *See, e.g.*, *United States v. Kriesel*, 508 F.3d 941, 950 (9th Cir. 2007) (upholding compulsory DNA sampling of non-violent felons); *United States v. Kincade*, 379 F.3d 813, 816, 840 (9th Cir. 2004) (en banc) (upholding compulsory DNA sampling of violent felons).

⁶⁸ Richard Willing, *With DNA Databases on Fast Track, Legal Questions Loom*, USA TODAY, Mar. 1, 1999, at 5A.

⁶⁹ FBI, CODIS-NDIS Statistics: Measuring Success, <http://www.fbi.gov/hq/lab/codis/clickmap.htm> (last visited Apr. 21, 2010) [hereinafter CODIS Measuring Success]. State and federal agencies added more than 2.8 million

contributor to this data pool, maintaining more than 1.25 million offender profiles as of February 2010—twice as many as Florida, the state with the second-largest offender database.⁷⁰ State lawmakers decide whom to include within each state's offender database.⁷¹ DNA collection laws initially targeted people convicted of a felony, or certain kinds of felonies, but some states have expanded offender databases to include arrestees and people convicted of misdemeanors.⁷² In early 2009, California began storing the DNA profiles of all adults arrested for a felony.⁷³

Law enforcement agencies search these databases to find investigative leads in criminal cases. As of February 2010, CODIS had produced more than 107,600 hits linking crime scenes through a common perpetrator, or linking a known offender to an unsolved crime.⁷⁴ Some of these cases, known as “cold hits,” had languished for years and would likely have remained unsolved if not for a match or partial match in CODIS.⁷⁵ The FBI and state law-enforcement agencies are becoming increasingly reliant on cold hits to solve

offender profiles between October 2007 and February 2010. *See* CODIS Measuring Success, *supra*; FBI, National DNA Index System, <http://www.fbi.gov/hq/lab/codis/national.htm> (last visited Apr. 16, 2010).

⁷⁰ CODIS Statistics, *supra* note 10.

⁷¹ DNA.gov, State DNA Database Statutes, <http://www.dna.gov/statutes-caselaw/state-statutes/database/> (last visited Apr. 21, 2010). Federal law governs DNA collection from people arrested, charged, or convicted for federal crimes. *See* 42 U.S.C. § 14135a(a)(1)(A) (West 2010).

⁷² Compare DNA.GOV, 2000 FORENSIC DNA EXPANSION LEGISLATION, available at <http://www.dnaresource.com/documents/2000DNADatabaseExpansionbills.pdf> (last visited Apr. 21, 2010), with DNA.GOV, 2009 DNA DATABASE LEGISLATION, available at <http://www.trendtrack.com/taxis/app/viewrpt?event=495bdbf6ba> (last visited Apr. 29, 2010), and DNA.GOV, 2008 DNA DATABASE EXPANSION LEGISLATION (Oct. 22, 2008), available at <http://www.dnaresource.com/documents/2008DNAExpansionLegislation.pdf>. For a year-by-year snapshot of state DNA database expansion legislation, visit <http://www.dnaresource.com/expansion.html>.

⁷³ Information Bulletin, Cal. Dep't of Justice (Dec. 15, 2008), available at http://ag.ca.gov/bfs/pdf/69IB_121508.pdf. Voters authorized the inclusion of arrestees in 2004. *Id.*

⁷⁴ CODIS Measuring Success, *supra* note 69.

⁷⁵ Some jurisdictions allow prosecutors to circumvent statutes of limitations by filing charges against suspects identified only by their DNA profiles. MICHAELIS ET AL., *supra* note 7, at 24–25.

crimes.⁷⁶ The FBI, which reported 19,500 cold hits by November 2004, had made more than 47,000 cold hits by March 2007.⁷⁷ It took the state of Virginia eight years to make 1000 cold hits, and eighteen months to make another thousand.⁷⁸

A cold hit in the database—the same kind of lead that pointed police to John Puckett—is sometimes the only evidence linking a person to a crime.⁷⁹ Some of the offenses in cold-hit cases are decades old, increasing the likelihood that witnesses have died or do not remember events with the clarity they once did.⁸⁰ The Puckett case is an example of this; by the time Puckett's case went to trial, Diana Sylvester's landlady and an early suspect were dead, and police had lost a bloody parking ticket they found in the suspect's car and matched to Sylvester's blood type.⁸¹ Furthermore, most courts have upheld convictions resting almost entirely on DNA evidence.⁸² This is true even when corroborating evidence is weak or conflicts with other evidence.⁸³ In these cases, jurors might be swayed by a form of the

⁷⁶ Yun S. Song et al., *Average Probability that a "Cold Hit" in a DNA Database Search Results in an Erroneous Attribution*, 54 J. FORENSIC SCI. 22, 22 (2009).

⁷⁷ *Id.*

⁷⁸ *Id.* Of the first 1000 cold hits in Virginia, 100 resulted in guilty pleas or guilty verdicts, seven resulted in not-guilty verdicts, and fifty-three were not prosecuted. *Id.* The remaining 752 were pending at the time of the survey. *Id.*

⁷⁹ See Murphy, *supra* note 37, at 738 (noting that DNA and other forensic technologies "allow for the identification of perpetrators even in the absence of any other evidence"). Professor Murphy notes that the lack of central record keeping, among other reasons, make it "difficult to determine the frequency with which the government presently brings cases in which the only evidence is genetic material." Murphy, *supra* note 37, at 742. Few laboratories monitor a case after the match is made, and they do not track whether investigators turned up evidence beyond the match in the database. *Id.* Also, DNA is daunting evidence likely to result in a guilty plea. *Id.*

⁸⁰ See Song et al., *supra* note 76, at 22. Federal funding programs encourage states to reopen unsolved cases and conduct DNA testing. *Id.* at 22, 24 n.7 (citing Justice for All Act of 2004, Pub. L. No. 108-405, 118 Stat. 2260).

⁸¹ Smith, *supra* note 4.

⁸² Song et al., *supra* note 76, at 22, 24 nn.8–10 (citing *Roberson v. State*, 16 S.W.3d 156, 170 (Tex. Crim. App. 2000); *People v. Johnson*, 139 Cal. App. 4th 1135 (Cal. Ct. App. 2006)).

⁸³ Song et al., *supra* note 76, at 23, 24 nn.11, 15 & 18 (citing *United States v. Jenkins*, 887 A.2d 1013 (D.C. 2005) (allowing a murder and burglary prosecution to proceed based largely on genetic evidence despite the fact that another man was found in possession of

CSI Effect, a phenomenon in which jurors inculcated by the television portrayal of foolproof forensic evidence come to expect it in all criminal cases, and when they get it, to trust it implicitly.⁸⁴ DNA evidence is especially persuasive to juries.⁸⁵ On average, jurors believe DNA evidence is 95 percent accurate and 94 percent persuasive of guilt, according to studies.⁸⁶

John Puckett is not the only accused whose DNA match is described by disparate statistics. During the proceedings of the National Commission on the Future of DNA Evidence in 2000, the chair, geneticist James Crow, mentioned news of a DNA match between crime scene evidence and a person who turned out to be innocent.⁸⁷ “The *New York Times* regarded that as surprising,” Crow said. “I regard it as not at all surprising.”⁸⁸ In an American court, the prosecution would likely have told the jury that the probability of a coincidental match between the innocent man and the true perpetrator was one in thirty-seven million.⁸⁹ However, when calculated to account for the fact that investigators made the match

the victim’s credit cards the day after the incident); *Riggs v. State*, 809 N.E.2d 322 (Ind. 2004) (non-DNA evidence: defendant lived near victim); *State v. Hunter*, 861 N.E.2d 898, 901 (Ohio Ct. App. 2006) (upholding a rape conviction based on a semen sample although “literally no other evidence” linked the defendant to the crime); *Convicted Murderer Seeks Retrial*, KALAMAZOO GAZETTE, May 10, 2006, at Local News (defendant convicted of 36-year-old murder despite the fact that DNA tests turned up his profile along with that of a person who was 4 years old at the time of the murder)).

⁸⁴ Prosecutors believe that the *CSI Effect* leads juries to acquit in the absence of scientific evidence, while defense attorneys say it explains jurors’ willingness to convict based upon forensic evidence alone. Michael Mann, *CSI Effect: Better Jurors Through Television and Science*, 24 BUFF. PUB. INT. L.J. 211, 213 (2005–2006). Yet Professors Simon Cole and Rachel Dioso-Villa examined acquittal rates before and after the show first aired, and did not find enough of a change to “warrant panic about the existence of a *CSI Effect*.” Simon A. Cole & Rachel Dioso-Villa, *Investigating the ‘CSI Effect’ Effect: Media and Litigation Crisis in Criminal Law*, 61 STAN. L. REV. 1335, 1364 (2009).

⁸⁵ Lieberman et al., *supra* note 1.

⁸⁶ *Id.* at 52–53.

⁸⁷ Proceedings, Nat’l Comm’n on the Future of DNA Evidence, Research and Dev. Working Group Report (Apr. 9, 2000), *available at* <http://www.ojp.usdoj.gov/nij/topics/forensics/events/dnamtgtrans9/trans-f.html>.

⁸⁸ *Id.*

⁸⁹ *Id.*

through a database search, the chance of a match grew to two percent.⁹⁰

Even without the complicating factor of a database search, DNA statistics rest upon assumptions. Population geneticists developed these statistics a decade ago using databases containing about 1000 people divided into discrete racial groups.⁹¹ These statistics assume that such a small database is adequate to estimate the rarity of DNA profiles in the general population.⁹² DNA rarity statistics also assume that the DNA markers within DNA profiles are independent of each other, rather than inherited together.⁹³ But if some strings of numbers within DNA profiles are dependent, meaning they are inherited together, certain DNA profiles might be less rare than DNA statistics say they are. Finally, DNA rarity statistics mathematically correct for the fact that people are related to one another and thus are more likely to share DNA patterns.⁹⁴ Large-scale empirical research would determine whether this correction is adequate.⁹⁵

Analysis of the database could also prove useful for other reasons. Access to the national DNA database would let scientists test the frequency with which three-person DNA mixtures are consistent with multiple people's DNA.⁹⁶ Further, studies of the database could shed light on the number of database errors that result from typos.⁹⁷ A quality-control study in Victoria, Australia, found that this kind of error occurred once every 300 entries.⁹⁸ These errors likely make

⁹⁰ *Id.* This figure—2 percent—is calculated by multiplying the random match probability (1 in 37 million) by 700,000, the number of profiles in the database. *Id.* For a description of this calculation, known as the NRC II formula, see *supra* note 8.

⁹¹ Mueller Decl. II, *supra* note 15, at 5; Mueller Decl. I, *supra* note 15, at 1; Letter from Krane et al., *supra* note 14, at 1.

⁹² Mueller Decl. I, *supra* note 15, at 1.

⁹³ *Id.*

⁹⁴ *Id.*; Letter from Dan E. Krane et al., *supra* note 14, at 2.

⁹⁵ Mueller Decl. I, *supra* note 15, at 1; Letter from Dan E. Krane et al., *supra* note 14, at 2 (noting that the national DNA database provides an opportunity for “real world tests of propositions that previously have been addressable only by simulation,” including the significance of relatives in the database).

⁹⁶ Letter from Dan E. Krane et al., *supra* note 14, at 2.

⁹⁷ *Id.* at 3.

⁹⁸ *Id.*

individuals within the database invisible to technicians searching for matches.⁹⁹

While they are not the large-scale analyses of offender databases that scientists say are warranted, technicians have searched states' offender databases for matches within them. In 2001, Arizona crime laboratory analyst Kathryn Troyer ignited interest in the study of offender databases when she found a nine-marker match between two men who did not appear related.¹⁰⁰ One was black and the other was white, and they had different names and birthdates.¹⁰¹ The Arizona database contained 10,000 profiles at the time, and DNA statistics put the chances of a coincidental match like the one Troyer found at one in 754 million for whites and one in 561 million for blacks.¹⁰² Four years later, a judge ordered the Arizona laboratory to conduct another search for DNA matches within the database.¹⁰³ By then its contents had swelled to 65,493 profiles, of which 122 pairs matched at nine markers and twenty pairs matched at ten markers.¹⁰⁴ No one has shown that the people associated with these profiles are related.¹⁰⁵

The significance of Troyer's results is a matter of debate. Some scientists say statistical principles explain the high number of matches within a small database, especially given the number of close relatives that tend to populate DNA databases.¹⁰⁶ And the nature of Troyer's search was different from the normal procedure, which is to compare

⁹⁹ *Id.*

¹⁰⁰ Smith, *supra* note 4, at 70.

¹⁰¹ *Id.*

¹⁰² *Id.*

¹⁰³ *Id.*

¹⁰⁴ *Id.*

¹⁰⁵ Jason Felch & Maura Dolan, *DNA: Genes as Evidence: FBI Resists Scrutiny of 'Matches'*, L.A. TIMES, July 20, 2008, at A1. The search turned up two more matches—one at eleven markers and one at twelve—that later turned out to be relatives. *Id.*

¹⁰⁶ Steven P. Myers, *Felon-To-Felon STR Partial Profile Matches in the Arizona Database: Don't Panic!*, available at http://www.nasams.org/forensics/for_lib/Documents/1148592247.61/Myers%20CAC%20Presentation.pdf (last visited Apr. 15, 2010); MICHAELIS ET AL., *supra* note 7, at 130 (noting that the statistical phenomenon at play is known as the Birthday Problem). For an explanation of the Birthday Problem, see Scott Simon, *Math Guy: The Birthday Problem*, NPR, Mar. 19, 2005, <http://www.npr.org/templates/story/story.php?storyId=4542341>.

a specific DNA profile to each profile within the database.¹⁰⁷ In Arizona in 2005, that would have involved making 65,493 comparisons—one for every profile in the database. Looking for matches among profiles within the database involved some two billion comparisons, making it much more likely that this kind of search would turn up matches.¹⁰⁸

Other statisticians believe that the Arizona matches make sense only if the database contained between 1000 and 3000 pairs of siblings.¹⁰⁹ In a declaration, Laurence D. Mueller, a professor of ecology and evolutionary biology at University of California, Irvine, wrote that the Arizona results “suggest that the [current] model may contain flaws.”¹¹⁰

Spurred by events in Arizona, defense attorneys began requesting similar searches of other states’ criminal-offender databases.¹¹¹ The FBI and state law enforcement agencies have balked, saying these searches would: (1) lead the FBI to expel states from the CODIS network; (2) violate convicted offenders’ privacy; (3) monopolize databases needed to catch criminals; and (4) be fruitless given the database’s unsuitability for study.¹¹² These concerns are explored in depth in the following part of this article.

California courts, and those of other states, have blocked “Arizona searches.”¹¹³ Courts in Illinois and Maryland have allowed them. Illinois’ 220,000-person database yielded 903 pairs of profiles that matched at nine or more markers.¹¹⁴ Maryland’s fewer-than-30,000-person database turned up thirty-two pairs of profiles that matched at nine or more markers.¹¹⁵ Of these, three pairs were perfect thirteen-

¹⁰⁷ Felch & Dolan, *supra* note 105.

¹⁰⁸ *Id.*

¹⁰⁹ Mueller, *supra* note 30, at 106.

¹¹⁰ Mueller Decl. I, *supra* note 15, at 1.

¹¹¹ Felch & Dolan, *supra* note 105; *See also* Felch & Dolan, *supra* note 4.

¹¹² Felch & Dolan, *supra* note 105.

¹¹³ *Id.*

¹¹⁴ *Id.* People v. Wright, Nos. 1-07-3106, 1-07-3464, 2010 WL 1194903, at *12 (Ill. App. Ct. Mar. 26, 2010). Illinois’s matches helped persuade an appeals court in *People v. Wright* to order a new trial for a defendant denied his own Arizona search. People v. Wright, Nos. 1-07-3106, 1-07-3464, 2010 WL 1194903, at *12 (Ill. App. Ct. Mar. 26, 2010).

marker matches.¹¹⁶ Population geneticists say these three pairs are probably duplicates, or they belong to identical twins or brothers.¹¹⁷ But if they are matches between unrelated people, they “defy[] odds as remote as 1 in 1 quadrillion.”¹¹⁸ No one knows which is true because “Maryland officials never did the research to find out.”¹¹⁹

Even if state officials had inquired further, Arizona searches are not the large-scale, rigorous studies statisticians and population geneticists say are necessary to test the assumptions underlying DNA statistics. Bruce Weir, a statistician at the University of Washington whose population genetics studies are the basis for forensic DNA probabilities, told the *Los Angeles Times* that offender databases could be used to test his estimates.¹²⁰ “Instead of saying we predict there will be a match, let’s open it up and look,” he said.¹²¹

Many of his colleagues agree. Science handed DNA technology to law enforcement, and scientists want to make sure their discoveries are still valid, especially given their use identifying suspects. “[S]tudies of model data sets have their place,” wrote sixteen professors in a letter to the FBI asking it to release the offender database, without names, for study.¹²² But “definitive results can be obtained only with real data.”¹²³ In a letter, the FBI recognized the need for “revalidating previous assumptions” and said it was “exploring ways to address” some of the points the scientists made.¹²⁴ Nonetheless, the agency denied their request days after a larger group of thirty-nine academics and lawyers published a letter to the editor of *Science* calling for scrutiny of DNA databases.¹²⁵ As they note, the FBI

¹¹⁵ Felch & Dolan, *supra* note 105.

¹¹⁶ *Id.*

¹¹⁷ *Id.*

¹¹⁸ *Id.*

¹¹⁹ *Id.*

¹²⁰ *Id.*

¹²¹ *Id.*

¹²² *Id.*

¹²³ *Id.*

¹²⁴ Letter from D. Christian Hassell, *supra* note 50, at 2.

¹²⁵ *Id.* at 1–2; Krane et al., *supra* note 15.

has published no research derived from the DNA information it holds, nor has it allowed others to do so.¹²⁶

State law enforcement agencies also resist this research. In a motion to quash a defense attorney's subpoena of the California database—a subpoena supported by non-government scientists from major research universities—lawyers for California's Department of Justice (DOJ) characterized the defense's interest in the contents of the offender database as a desire to "conduct mysterious research or engage in data manipulation."¹²⁷ Yet peer-reviewed scientific inquiry is the antidote to mystery and manipulation. It is a collaborative process in which scientists make discoveries, publish those discoveries in peer-reviewed journals, test them some more, and let the larger scientific community reach consensus as to the quality of the original discovery.¹²⁸ Only more research will reveal whether the science underlying forensic DNA databases "has been performed in accordance with established standards in the field," wrote Montgomery Slatkin, a population geneticist at University of California, Berkeley, in a declaration in support of the database subpoena. "No such confidence exists at the present time," he said.¹²⁹

Other scientists agree with Slatkin.¹³⁰ Research on the offender database "is long overdue," said Keith Devlin, a mathematics professor at Stanford University.¹³¹ Given that the contents of a DNA database could be encoded to ensure anonymity, "it is . . . scientifically imperative that it be analyzed, [regardless] of any other considerations,"¹³² he said. "To do otherwise would run counter to the

¹²⁶ Krane et al., *supra* note 47.

¹²⁷ Ford Motion to Quash, *supra* note 32, at 15.

¹²⁸ Decl. of Montgomery Slatkin, *supra* note 15 (arguing that "these databases should be made available to qualified researchers who are outside the forensic genetics community and the result of their analyses should be published in the scientific literature in accordance with the standard practice"); Mueller Decl. I, *supra* note 15, at 3 (noting that "this is how science works: through an open exchange of ideas and research, and the development of real world tests for existing hypotheses and theories").

¹²⁹ Decl. of Montgomery Slatkin, *supra* note 128.

¹³⁰ See, e.g., Letter from Dan E. Krane et al., *supra* note 14.

¹³¹ Letter from Keith Devlin, *supra* note 15.

¹³² *Id.*

very principles of sound science and of sound legal practice based on proven science.”¹³³

The idea that longstanding scientific principles require continued study can be a troublesome concept to judges, who seek finality in the law. Once a legal issue has been decided, its power as precedent can outlive its wisdom in light of new information.¹³⁴ At a 2006 hearing to decide whether to grant a defense attorney’s subpoena of the California database, the trial judge asked Professor Slatkin if the statistical analysis he advocated was likely to reveal flaws in the way DNA statistics are currently calculated.¹³⁵ “I had to say I didn’t know,” Slatkin said.¹³⁶ “I said answering the question was the point of the research.”¹³⁷ The judge denied the database search.¹³⁸

In a rare departure,¹³⁹ a three-judge panel in Illinois took the view that denying a defendant access to an offender DNA database is an error serious enough to warrant a new trial.¹⁴⁰ The defendant, Harvey Wright, was identified through a cold hit and prosecuted for the 1998 rape of a 15-year-old girl.¹⁴¹ The victim did not recognize Wright, and the only physical evidence tying him to the crime was that his DNA profile shared nine markers with DNA from the victim’s rape kit.¹⁴²

¹³³ *Id.* (emphasizing that “[a]dditional information, in particular empirical statistical information about the very data used in criminal investigations, can surely not do otherwise than make our judicial system even more reliable than it is already. I am unaware of any creditable reason to do otherwise”).

¹³⁴ For example, courts continue to admit fingerprints and other forensic evidence despite the fact that these methods have no demonstrated scientific basis. *See NAS REPORT, supra* note 54. A trial court that questioned the probative value of fingerprint evidence was ridiculed to the point of reversing itself. Murphy, *supra* note 37, at 768 n.205.

¹³⁵ Jon Jefferson, *Cold Hits Meet Cold Facts: Are DNA Matches Infallible?*, TRANSCRIPT, Spring 2008, at 29, 33.

¹³⁶ *Id.*

¹³⁷ *Id.*

¹³⁸ *Id.*

¹³⁹ *See supra* note 32.

¹⁴⁰ *People v. Wright*, Nos. 1-07-3106, 1-07-3464, 2010 WL 1194903 (Ill. App. Ct. Mar. 26, 2010).

¹⁴¹ *Id.* at *1, *4–*6.

¹⁴² *Id.* at *1, *6.

The jury heard infinitesimal rarity statistics—one in 420 trillion for African Americans, one in 670 trillion for whites, or one in 2.9 quadrillion for Hispanics.¹⁴³ In the hopes of undermining these numbers, the defendant asked the trial court for permission to search the Illinois's offender database for nine-marker matches within it.¹⁴⁴ The trial court refused, and in 2007 Wright was convicted and sentenced to life in prison.¹⁴⁵ The appellate panel held that denying his motion was an “error so serious that it affected the fairness of the defendant’s trial and challenged the integrity of the judicial process.”¹⁴⁶ It noted Arizona search results in Illinois and elsewhere, and said that the trial court’s decision barred the “defendant’s access to evidence that could have assisted him in establishing his innocence” by “casting a serious doubt” upon the DNA evidence presented by the prosecution.¹⁴⁷

If most courts have been reluctant to address this problem,¹⁴⁸ so has the executive branch. As noted above, the FBI rejected a request by sixteen scientists seeking access to the contents of the national offender database.¹⁴⁹ It did so despite the Obama Administration’s warning that the “public must trust the science and scientific process informing public policy decisions.”¹⁵⁰ Even though “science and technological information . . . developed and used by the Federal

¹⁴³ *Id.* at *6.

¹⁴⁴ *Id.* at *2.

¹⁴⁵ *Id.* at *1, *3, *7.

¹⁴⁶ *Id.* at *11.

¹⁴⁷ *Id.* at *11–*12. This ruling might have been influenced by the fact that Illinois is among the few states with a statute allowing defendants to request searches of the state’s offender database. Murphy, *supra* note 37 at 790–91 (citing 725 ILL. COMP. STAT. § 5/116-5 (2005) and GA. CODE ANN. § 24-4-63 (2005)). However, defendants’ “rights to due process and the assistance of counsel clearly contemplate” this kind of search even without a statutory framework. *Id.*

¹⁴⁸ *See supra* note 32.

¹⁴⁹ Letter from D. Christian Hassell, *supra* note 50.

¹⁵⁰ Press Release, The White House, Memorandum for the Heads of Executive Departments and Agencies (Mar. 9, 2009), *available at* http://www.whitehouse.gov/the_press_office/Memorandum-for-the-Heads-of-Executive-Departments-and-Agencies-3-9-09.

Government . . . should ordinarily be made available to the public," the FBI did not heed.¹⁵¹

Authorities in Australia have taken a different approach. In 2003, technicians there matched DNA from the clothing of a murdered child to a rape victim with no connection to the case—a match that prevailing science would say has a one in 269 million chance of being coincidental.¹⁵² During the inquest that followed, the Victoria Police Forensic Services Centre made its 15,000-profile offender database available to an American geneticist, who studied it for partial matches.¹⁵³ The database was too small to determine whether forensic DNA matches are as common as current statistics say they are.¹⁵⁴

A court system that leans so heavily on science in the administration of justice must also embrace the scientific tenet that more knowledge is better than less knowledge. This idea likely drives those who believe that current DNA statistics are accurate, yet see value in double and triple checking. "I, for one, would be surprised if major modifications were required," wrote David Kaye, a law professor at Arizona State University with an advanced degree in science.¹⁵⁵ "But expectations are no substitute for inquiry."¹⁵⁶ The next part of this article evaluates government agencies' arguments against independent research of forensic DNA databases.

II. PRO-SECREC Y ARGUMENTS DO NOT HOLD UP

As noted in Part I, the FBI and other law enforcement agencies protect the secrecy of forensic DNA profiles. Granting outsiders access to them would, they say: (1) lead the FBI to expel states from the CODIS network; (2) violate convicted offenders' privacy; (3) monopolize databases needed to catch criminals; and (4) prove

¹⁵¹ *Id.*

¹⁵² Gary Tippet & Selma Milovanovic, *Who Killed Jaidyn? A Sad Saga Continues...*, THE AGE (MELBOURNE, AUSTL.), Nov. 18, 2003, available at <http://www.theage.com.au/articles/2003/11/17/1069027047383.html>; WILLIAM C. THOMPSON, VICTORIA STATE CORONER'S INQUEST INTO DEATH OF JAIDYN LESKIE 4 (Dec. 3, 2003), available at <http://www.bioforensics.com/articles/Thompsonreport.pdf>.

¹⁵³ Mueller, *supra* note 30, at 102.

¹⁵⁴ *Id.*

¹⁵⁵ Letter from David Kaye, *People v. Ford*, No. SCN 2151192 (Cal. Super. Ct. Oct. 3, 2008).

¹⁵⁶ *Id.*

fruitless because offender databases are unsuitable for testing the reliability of DNA statistics. This part will analyze these arguments in turn.

A. EXPULSION FROM CODIS: AN EMPTY THREAT

Government lawyers continue to argue that nothing short of total secrecy protects a state's membership in the national CODIS system.¹⁵⁷ Their concern is that releasing a state's offender database to an outside researcher would allow the FBI to expel the state from the national DNA network, cutting off that state's law enforcement agencies from offender databases around the country. Government lawyers in California have repeatedly raised this argument in their efforts to prevent defense experts from accessing the state's offender database.¹⁵⁸ They cite two sources of support for their position—the federal CODIS statute and California's Memorandum of Understanding with the FBI.¹⁵⁹ As this article will show in Part III, the CODIS statute expressly authorizes study of offender databases, and the FBI cannot expel a state for doing what the statute allows. The Memorandum of Understanding between California and the FBI merely requires California to “take reasonable precautions to prevent unauthorized persons from accessing the CODIS software.”¹⁶⁰ Access that is legal under the CODIS statute is likely authorized under the Memorandum of Understanding.

The threat of expulsion from CODIS has proved to be an empty one.¹⁶¹ An investigation by the *Los Angeles Times* found that the FBI's director of CODIS “advised state officials to raise the risk of expulsion with a judge but told the officials that expulsion was unlikely to actually happen.”¹⁶² As noted above, the CODIS statute allows court-ordered database access. Furthermore, the FBI and other states need California's more than one million DNA profiles at least as much as

¹⁵⁷ Felch & Dolan, *supra* note 105.

¹⁵⁸ Ford Motion to Quash, *supra* note 32, at 9; Motion to Quash Subpoena Duces Tecum, *People v. Davis*, No. SCN 190226, 8 (Cal. Super. Ct. Feb. 3, 2006) [hereinafter *Davis* Motion to Quash].

¹⁵⁹ Ford Motion to Quash, *supra* note 32, at 9; *Davis* Motion to Quash, *supra* note 158, at 8.

¹⁶⁰ *Davis* Motion to Quash, *supra* note 158, at 8.

¹⁶¹ Felch & Dolan, *supra* note 105.

¹⁶² *Id.*

California needs the rest of the CODIS network;¹⁶³ the FBI had uploaded only 156,344 profiles to CODIS as of February 2010.¹⁶⁴

To prevent a court from ordering an Arizona search of Illinois's offender database, Thomas Callaghan, director of the FBI's CODIS system, met with state lab officials in 2006 about "how to fight this."¹⁶⁵ According to the lab officials' summary of the meeting, Callaghan told them to tell the court that Illinois could be disconnected from CODIS if it authorized an Arizona search.¹⁶⁶ He then went on to assure lab officials that "it would in fact be unlikely that [Illinois] would be disconnected."¹⁶⁷ Callaghan disputes this account of the meeting, but acknowledged that "it takes a lot for a state to be cut off from the national database."¹⁶⁸

State and federal officials used the same strategy in Maryland.¹⁶⁹ There, a state lawyer argued that conducting an Arizona search of the Maryland database would be "catastrophic" because it could lead to the state's expulsion from CODIS.¹⁷⁰ The state's DNA administrator, Michelle Groves, filed a supporting affidavit edited by FBI officials.¹⁷¹ Before submitting the affidavit to the court, Groves sent an e-mail to FBI officials saying, "Let's see if this will work."¹⁷²

It did not work. The trial court held firm.¹⁷³ "This court will not accept the notion that the extent of a person's due process rights hinges solely on whether some employee of the FBI chooses to

¹⁶³ *Id.*

¹⁶⁴ CODIS Statistics, *supra* note 10.

¹⁶⁵ Felch & Dolan, *supra* note 105 (referring to a crime lab's conversation log dated July 5, 2006, on file with author).

¹⁶⁶ *Id.*

¹⁶⁷ *Id.*

¹⁶⁸ *Id.*

¹⁶⁹ *Id.*

¹⁷⁰ *Id.*

¹⁷¹ *Id.*

¹⁷² *Id.*

¹⁷³ *Id.*; State v. Davis, No. CT 05-0024X (Md. Cir. Ct. Aug. 4, 2006) (Platt, J.).

authorize the use of the [database] software,” the court wrote.¹⁷⁴ Searches went ahead in Illinois and Maryland.¹⁷⁵ Both are still members of CODIS.¹⁷⁶

Despite this history, government lawyers continue to invoke the threat of expulsion from CODIS. Arguing successfully against a subpoena of California’s offender DNA database, state lawyers warned of “expulsion from the national CODIS network.”¹⁷⁷ In reality, the FBI is either unwilling to expel a needed state from its DNA network, or searches and studies of offender databases do not violate federal law. This article will explore the latter point in Part III.

B. EXAMINATION OF DNA DATABASES DOES NOT THREATEN PRIVACY

Government agencies opposed to research involving DNA databases invoke the privacy of people whose DNA profiles would be studied.¹⁷⁸ In a motion to quash a subpoena of California’s offender database, the state DOJ argued that secrecy is “a significant factor in maintaining the program’s constitutionality.”¹⁷⁹ It cited state and federal cases in which courts have relied upon the limited use of criminal offender DNA databases in holding them constitutional under the Fourth Amendment.¹⁸⁰ It specifically noted *United States v. Kincade*, where the Ninth Circuit relied upon statutory confidentiality protections to counter the argument that “soon, if not already, scientists will request access to what would serve as [a] preexisting goldmine of DNA data for their research.”¹⁸¹ The FBI has long sought to quell the fear that private companies might someday make use of

¹⁷⁴ Felch & Dolan, *supra* note 105; *State v. Davis*, No. CT 05-0024X (Md. Cir. Ct. Aug. 4, 2006) (Platt, J.).

¹⁷⁵ Felch & Dolan, *supra* note 105.

¹⁷⁶ CODIS Statistics, *supra* note 10.

¹⁷⁷ Ford Motion to Quash, *supra* note 32, at 9.

¹⁷⁸ Felch & Dolan, *supra* note 105.

¹⁷⁹ Ford Motion to Quash, *supra* note 32, at 7.

¹⁸⁰ *Id.* (citing *Alfaro v. Terhune*, 98 Cal. App. 4th at 492, 507-8 (Cal. Ct. App. 2002); *People v. King*, 82 Cal. App. 4th 1375, 1377 n.6 (Cal. Ct. App. 2000)).

¹⁸¹ *Id.* (citing *United States v. Kincade*, 379 F.3d 813, 837 n.33 (9th Cir. 2004)).

offender DNA databases.¹⁸² “People hear government, DNA, computers, FBI, and they get wary,” said Stephen Niezgoda, CODIS program director.¹⁸³ “I understand that. But the system is built to maintain privacy.”¹⁸⁴ Privacy is incompatible with outside research, according to the FBI. Bruce Budowle, former director of the FBI’s DNA laboratory at Quantico, Virginia, warned that a breach of privacy and confidentiality would take place if the offender database were subject to statistical study.¹⁸⁵ “The names of individuals with matching and partial matching profiles would have to be disclosed to scientists and police when there is no criminal investigation underway,” he wrote.¹⁸⁶ “The names would be obtained because of a ‘research experiment.’”¹⁸⁷

CODIS does not contain offenders’ names,¹⁸⁸ and to the extent names are needed to distinguish among individuals, they can be encoded. Furthermore, this article does not propose that law enforcement agencies release offenders’ DNA profiles for general genetic research. Rather it argues that they should use DNA databases to answer questions relating directly to the government’s use of the database and DNA evidence.¹⁸⁹ Research into possible genetic roots of criminality or the study of offender profiles or samples for commercial gain would be inappropriate. But if law-enforcement agencies want to continue using DNA evidence in court, they must also be willing to investigate whether the statistics that accompany that evidence are accurate.

¹⁸² Finn, *supra* note 66.

¹⁸³ *Id.*

¹⁸⁴ *Id.*

¹⁸⁵ Bruce Budowle et al., *supra* note 43.

¹⁸⁶ *Id.*

¹⁸⁷ *Id.* Budowle has used offender profiles in his own research. See *infra* note 231 and surrounding text.

¹⁸⁸ See *infra* note 199.

¹⁸⁹ Releasing DNA databases for quality control supports the core purpose of forensic DNA collection—accurately identifying suspects—while other uses do not. See Lazer & Mayer-Schonberger, *infra* note 373.

Some geneticists question whether DNA profiles can ever truly be anonymous.¹⁹⁰ In August 2008, privacy concerns prompted the National Institutes of Health to remove from its public website the genetic information of some 60,000 patients who donated their DNA for research purposes on the condition that they not be linked to it.¹⁹¹ However, making genetic information available for download online, or incorporating it into patients' medical records, is a more radical dissemination than entrusting offenders' DNA profiles to a small number of scientists. The offender database contains the same kind of information already widely accessible through population-genetics databases that scientists used to develop current DNA statistics. California DOJ argued in a recent case that independent access to the offender database was unnecessary given the availability of these smaller DNA databases.¹⁹² It listed several of them, the largest of which contains 17,000 thirteen-marker profiles.¹⁹³ Since advocates can be trusted to handle these DNA profiles, surely statisticians and population geneticists can be trusted with criminal offender databases. In fact, some scientists have used offenders' DNA for scientific study, and in one of them, researchers published Illinois offenders' profiles, without names, in the *Journal of Forensic Science*.¹⁹⁴ Moreover, the fact that state law allows independent researchers to study newborns' DNA samples—a phenomenon explored in Part III—argues against privacy claims about offenders' DNA profiles.

¹⁹⁰ See B. Malin & L. Sweeney, *How (Not) to Protect Genomic Data Privacy in a Distributed Network: Using Trail Re-identification to Evaluate and Design Anonymity Protection Systems*, 37 J. BIOMEDICAL INFORMATICS 179 (2003). Malin and Sweeney examine the susceptibility of encrypted genetic information released in the healthcare context to be re-identified to named persons using publicly available records and patient-location visit patterns. *Id.* at 179. These kinds of re-identifications have taken place in other contexts, including Google searches and Netflix queues. Posting of Tim Lee to Techdirt, <http://techdirt.com/articles/20071130/114005.shtml> (Nov. 30, 2007, 14:41 PST).

¹⁹¹ Jason Felch, *DNA Databases Blocked from the Public*, L.A. TIMES, Aug. 29, 2008, at A31.

¹⁹² Ford Motion to Quash, *supra* note 32, at 15.

¹⁹³ *Id.*

¹⁹⁴ Mueller, *supra* note 30, at 101 (referring to W.E. Frank et al., *Y Chromosome STR Haplotypes and Allele Frequencies in Illinois Caucasian, African American, and Hispanic Males*, 51 J. FORENSIC SCI. 1207 (2006)); Bruce Budowle et al., *Population Studies on Three Native Alaska Population Groups Using STR Loci*, 129 FORENSIC SCI. INT'L 51 (2001).

Scientists interested in DNA database research acknowledge that offender profiles should not be publicly available.¹⁹⁵ However, these profiles can be used for research and remain confidential. Professor Keith Devlin says “it would be a fairly straightforward task to create an anonymized copy of an existing offender DNA database.”¹⁹⁶ Scientists are accustomed to working with sensitive data, and confidentiality is as assured here as it is in all genetic studies involving human subjects.¹⁹⁷

Furthermore, if the constitutionality of DNA databases rests in part upon their confidentiality, it also rests upon the idea that they contain useless genetic information.¹⁹⁸ That is, that convicted offender databases contain DNA sequences believed to reveal very little, if anything, about convicted offenders. Law enforcement officials have emphasized this since CODIS began.¹⁹⁹

¹⁹⁵ Decl. of Montgomery Slatkin, *supra* note 128. Professor David Kaye believes “that the database should be available for statistical research by the entire scientific community.” Letter from David Kaye, *supra* note 155. He notes that the information “can and should be released in an anonymized form so as not to compromise the privacy of the sources of the data.” *Id.*

¹⁹⁶ Letter from Keith Devlin, *supra* note 15.

¹⁹⁷ Decl. of Montgomery Slatkin, *supra* note 128. Furthermore, scientists entrusted with the database in the context of a criminal case would surely “honor all court orders, including protective orders, and would not use the data for any unlawful or inappropriate purpose.” Mueller Declaration II, *supra* note 15, at 3. “To suggest otherwise is misinformed and misguided.” *Id.* Mueller is one of sixteen scientists who asked the FBI for access to the national offender DNA database, minus identifying information. Letter from Dan E. Krane et al., *supra* note 14, at 1. In their request, they offered to sign “a reasonable confidentiality agreement.” *Id.* at 3.

¹⁹⁸ See, e.g., *United States v. Weikert*, 504 F.3d 1, 13 (1st Cir. 2007) (“the possibility that junk DNA may not be junk DNA some day [sic] ... does not significantly augment Weikert's privacy interest in the present case”); *United States v. Kincade*, 379 F.3d 813, 837 (9th Cir. 2004) (emphasizing that a DNA profile “establishes only a record of the defendant's identity”).

¹⁹⁹ The unrevealing nature of forensic DNA has been central to legal arguments in defense of its collection. In *United States v. Kincade*, a case that challenged the constitutionality of forcing offenders to surrender DNA samples, the federal government argued that the “genetic markers contained in the DNA sample serve as a ‘genetic fingerprint’ in that they uniquely identify an individual, but do not convey any other information about the person, such as physical or medical characteristics.” Appellee's Brief, *United States v. Kincade*, 2002 WL 32181457, at *8–*9 (9th Cir. Oct. 30, 2002). It noted that DNA records in CODIS do not contain names. *Id.* at 9. They contain: “(1) an agency identifier for the agency submitting the DNA profile; (2) the specimen identification number; (3) the DNA profile; and (4) the name of the personnel associated with the DNA analysis.” *Id.*

In 1999, an FBI spokesperson told *USA Today* that the type of DNA used for forensic identification reveals no genetic information.²⁰⁰ “Our lab people call it ‘junk DNA,’” he said.²⁰¹ Barry Duceman, director of biological sciences at the New York State Police Forensic Investigation Center, made similar assurances.²⁰² “It’s junk DNA, something that doesn’t even carry characteristics that could identify a person, even height or eye color,” he told a newspaper in 2000.²⁰³ California officials have also minimized the privacy concerns involved in collecting offenders’ DNA profiles. “California samples include only the most basic genetic materials, sometimes referred to as ‘junk DNA,’” said Hallye Jorden, a spokesperson for former state Attorney General Bill Lockyer, in 2004.²⁰⁴ “All these samples have are identifying markers about who these people are,” she said.²⁰⁵

Even so, insurance companies or employers might use offender DNA databases to screen patients or job applicants. Laws exist to prevent this kind of discrimination,²⁰⁶ but to guard against it states could require that law enforcement encrypt the offender database before releasing it to outsiders. Scientists want to know the rarity of certain combinations of genetic code. While the common convention is to represent genetic code as strings of universally accepted

²⁰⁰ Willing, *supra* note 68.

²⁰¹ *Id.*

²⁰² Linda Trischitta, *As DNA Use Rises, so Do Concerns*, TIMES UNION (ALBANY, N.Y.), Mar. 5, 2000, at A1.

²⁰³ *Id.* This same point came up in a television interview with Angelo Della Manna, head of DNA analysis for the state of Alabama. *60 Minutes: A Not So Perfect Match* (CBS television broadcast July 15, 2007) (transcript on file with author). The correspondent asked Manna, “When you look at [a DNA profile], can you tell me what color eyes that person has, or how tall they are? Or anything like that?” Manna responded: “No, not at all. The areas that we look at are commonly referred to as ... junk DNA.” *Id.*

²⁰⁴ Matt Krupnick, *Future to Grow More Clear on DNA Sampling; Ballot Measure, Trials and Supreme Court Decision Will Affect Application of Technology*, CONTRA COSTA TIMES, Sept. 5, 2004, at A33.

²⁰⁵ *Id.*

²⁰⁶ See, e.g., Genetic Nondiscrimination Act of 2007, Pub. L. No. 110-233 (codified as amended in scattered sections of 42 U.S.C.). Professor Laurence Mueller notes that much of what DNA databases might reveal is publicly available. Mueller Decl. II, *supra* note 15, at 3. It would be “far easier and cheaper to simply check public records to see if someone had been convicted of [a] crime that qualified them for inclusion in these databases.” *Id.*

numbers, government agencies could easily encrypt them using a new set of letters, numbers, or even symbols. And to ward off concern about duplicate profiles in the database or the release of offenders' names, states could encrypt individuals' identifying information as well.

C. STATISTICAL DATABASE RESEARCH DOES NOT PREVENT POLICE FROM CATCHING CRIMINALS

Some opponents of offender database research contend that it would have "dire consequences," including corrupting the database or taking it away from its primary function of catching criminals.²⁰⁷ This argument is discredited, and Michael Chamberlain, the attorney who leads California DOJ's resistance to database disclosure, said his office had abandoned it.²⁰⁸

Nonetheless, practical concerns have been a central tenet in favor of secrecy. During litigation to avoid an Arizona search of the Maryland database, the state DNA administrator argued in court that such a search might corrupt the database system.²⁰⁹ The trial court ordered the search anyway.²¹⁰ The system did not crash.²¹¹

In *United States v. Berger*, a criminal DNA case in Washington, D.C., federal prosecutors said that performing an Arizona search on the national database could take as long as six months.²¹² "This seems to be a tremendously inflated estimate," wrote Professor Laurence Mueller in a declaration to the court.²¹³ It could be done in fewer than eight days, he estimated.²¹⁴ Furthermore, given that the entire CODIS database—which at that time contained more than 3.5 million profiles—could fit easily on a compact disk, the FBI need not interrupt

²⁰⁷ Felch & Dolan, *supra* note 105.

²⁰⁸ Telephone Interview with Michael Chamberlain, Deputy Attorney Gen., Cal. Dep't of Justice (Feb. 26, 2009).

²⁰⁹ Felch & Dolan, *supra* note 105.

²¹⁰ *Id.*; State v. Davis, No. CT 05-0024X (Md. Cir. Ct. Aug. 4, 2006) (Platt, J.).

²¹¹ Felch & Dolan, *supra* note 105.

²¹² Mueller Decl. II, *supra* note 15, at 3.

²¹³ *Id.*

²¹⁴ *Id.*

its investigatory work to do the analysis, he wrote.²¹⁵ The judge denied the search.²¹⁶

Professor Montgomery Slatkin also rejects the idea that statistical research of offender profiles might thwart the database's effectiveness as a law enforcement tool. The state could analyze the database using a backup data set, leaving CODIS to the task of fighting crime, he said.²¹⁷ "I back up my computer every night," Slatkin said. "I would hope the government does the same."²¹⁸

Laws in Illinois and other states allow criminal defendants and people convicted of crimes to ask law enforcement to conduct DNA database searches that might produce exculpatory evidence or exonerate them.²¹⁹ These searches presumably take the database away from its task of finding new cold hits, with no disastrous results. Even so, post-conviction DNA testing encountered the same kinds of warnings now raised in opposition to the study of DNA databases. In an editorial arguing in favor of DNA testing to exonerate people wrongly convicted, former FBI director William S. Sessions wrote: "There are always reasons—time, money, bureaucratic obstacles—that something cannot be done. But when it comes to justice and fairness, those reasons are just excuses."²²⁰

The excuse here—that outside researchers will corrupt or shut down DNA databases—is similarly without merit.

²¹⁵ *Id.*

²¹⁶ Email from Katherine Philpott, Special Litigation and Forensics Fellow, Public Defender Serv. for the Dist. of Columbia, to author (May 2, 2010, 22:11 PST) (on file with author).

²¹⁷ Jefferson, *supra* note 135, at 33.

²¹⁸ *Id.*

²¹⁹ Murphy, *supra* note 37, at 790–91 (citing 725 ILL. COMP. STAT. § 5/116-5 (2005) and GA. CODE ANN. § 24-4-63 (2005)); DNA.gov, The Impact of DNA Testing, DNA Initiative, <http://www.dna.gov/postconviction/handling-requests/legalissues/impact-of-dna-testing> (last visited Apr. 16, 2010) (citing N.Y. CRIM. PROC. LAW § 440.30(1-a) (McKinney Supp. 1999); 725 ILL. COMP. STAT. 5/116-3(a) (West Supp. 1998)).

²²⁰ William S. Sessions, *DNA Tests Can Free the Innocent. How Can We Ignore That?*, WASH. POST, Sept. 21, 2003, at B2.

D. OFFENDER DNA DATABASES ARE SUITABLE FOR STUDY

Government scientists have argued that criminal offender databases cannot be studied.²²¹ Geneticists and statisticians typically work with DNA from discrete racial and ethnic categories, while DNA databases contain profiles from diverse populations.²²² FBI scientists point out that offender databases hold the DNA of African Americans, Asians, Caucasians, Hispanics, Native Americans, and Oceanians, and “it is likely that the proportions of these groups in the database are not the same as they are in the greater US population.”²²³ The databases also contain duplicate profiles and very close relatives, which would tend to skew statistical results.²²⁴ Such factors make CODIS “an extremely poor database to analyze” for the purpose of testing DNA rarity assumptions, according to the FBI.²²⁵

While the FBI points out legitimate complicating factors, statistical methods can correct for them.²²⁶ These methods allowed researchers to examine offender profiles from Virginia and North Carolina to estimate the number of DNA markers necessary to identify a single individual through a database search.²²⁷ Scientists used samples from Illinois’s offender population to create a database of DNA profiles, and published those profiles in the *Journal of Forensic*

²²¹ Budowle et al., *supra* note 43, at 60.

²²² *Id.*

²²³ *Id.*

²²⁴ *Id.*

²²⁵ *Id.*

²²⁶ Decl. of Montgomery Slatkin, *supra* note 128. See also Letter from Keith Devlin, *supra* note 29 (noting that the lack of randomness and other properties of DNA databases “in no way pose[] a major obstacle to an analysis” and “reliable statistical techniques can be used to make allowances for the peculiarities of the population represented in the database”); Letter from David Kaye, *supra* note 155 (arguing that “although these databases contain unknown numbers of close relatives, which complicates the analysis, they have considerable statistical power because of their size”); Mueller Decl. I, *supra* note 15 (arguing that “because of its size, quality and the availability of records on each sample, [the national offender database] is well suited to empirical study”).

²²⁷ Mueller, *supra* note 30, at 101.

Science.²²⁸ Professor Laurence Mueller studied the Arizona matches, using statistics to figure out whether factors other than incorrect assumptions could account for them.²²⁹ Australian authorities opened a local offender database to a scientist's scrutiny after finding a DNA match that seemed to defy prevailing rarity statistics.²³⁰

Moreover, some of the same FBI scientists who say database profiles are unsuitable for scientific research have themselves used database profiles in their research.²³¹ Bruce Budowle, who once headed the FBI's laboratory division at Quantico, Virginia, and is now a professor at the University of North Texas, has consistently argued against the scientific study of genetic information collected from criminal offenders.²³² Even so, he has used offenders' DNA samples for his own research. In a study investigating the rarity of DNA from three Native Alaskan populations, he and other FBI scientists wrote that "the majority of the samples for this study were collected from convicted offenders required to provide a DNA sample by Alaskan Statute."²³³ To identify these offenders by ethnic affiliation, Budowle and his team relied on the information provided by the offenders at the time their samples were collected.²³⁴

Budowle and his co-authors are some of the same scientists who emphasized the "privacy and confidentiality" of offender DNA profiles in a 2009 paper opposing scientific research of this data set.²³⁵ While research conducted by government scientists might pose less of a privacy risk than that of outsiders, Budowle and his colleagues have also claimed that offenders' DNA profiles can only yield "flawed" and

²²⁸ *Id.* (referring to W.E. Frank et al., *Y Chromosome STR Haplotypes and Allele Frequencies in Illinois Caucasian, African American, and Hispanic Males*, 51 J. FORENSIC SCI. 1207 (2006)).

²²⁹ *Id.*

²³⁰ See *supra* note 98 and accompanying text.

²³¹ See Bruce Budowle et al., *supra* note 194.

²³² See, e.g., Budowle et al., *supra* note 43, at 60; Ford Motion to Quash, *supra* note 32 (appended Clarification of Statistical Issues Related to the Operation of CODIS); Editors & Editorial Board, Forensic Science Communications, <http://www.fbi.gov/hq/lab/fsc/current/editors.htm> (last visited Apr. 15, 2010).

²³³ Budowle et al., *supra* note 194, at 52.

²³⁴ See *id.*

²³⁵ Budowle, *supra* note 43, at 60.

“misleading” results.²³⁶ Budowle and his co-author, Ranajit Chakraborty, did not return calls and e-mails asking them to explain why offender DNA profiles were nonetheless suitable for their Alaska study.²³⁷

Other scientists say further study of offender databases could answer persistent questions about DNA statistics. In a declaration in support of an unsuccessful subpoena of the California database, Professor Slatkin said that despite their heterogeneity, databases are large enough that “great statistical power could be achieved from a detailed analysis of them.”²³⁸ Even California’s more than one million profiles—a fraction of the profiles in CODIS—could help test critical assumptions about the rarity of DNA profiles, he said.²³⁹ A study of the California DNA database would give researchers the benefit of working with the data actually used to identify suspects, eliminating the current dependence on disputed assumptions.²⁴⁰ As Keith Devlin, a Stanford University mathematician, asked, “Why argue about *what might be* when we can know for sure *what is*?”²⁴¹

The following part takes a closer look at DNA collection statutes and finds that they do not forbid independent research.

²³⁶ *Id.*

²³⁷ E-mail from author to Bruce Budowle, Professor, Univ. of N. Texas (Mar. 2, 2010 14:32 PST); E-mail from author to Bruce Budowle, Dir., Lab. Div., FBI, Quantico, Va. (Mar. 16, 2009 13:09 PST); E-mail from author to Ranajit Chakraborty, Professor, Ctr. for Genome Info., Univ. of Cincinnati (Mar. 16, 2009 13:09 PST); Voice Mail Message to Bruce Budowle, Dir., Lab. Div., FBI, Quantico, Va. (Feb. 27, 2009); E-mail from author to Bruce Budowle, Dir., Lab. Div., FBI, Quantico, Va. (Feb. 27, 2009 11:12 PST); Voice Mail Message to Ranajit Chakraborty, Professor, Ctr. for Genome Info., Univ. of Cincinnati (Feb. 27, 2009); E-mail from author to Ranajit Chakraborty, Professor, Ctr. for Genome Info., Univ. of Cincinnati (Feb. 27, 2009 11:22 PST).

²³⁸ Decl. of Montgomery Slatkin, *supra* note 128.

²³⁹ *Id.* Publicly available genetic databases are much smaller than offender databases, and thus less fruitful for study. Mueller Decl. II, *supra* note 15, at 2–3 (pointing out that a 17,000-person database, the largest publicly available, is “relatively small”).

²⁴⁰ Letter from Prof. Keith Devlin, *supra* note 29.

²⁴¹ *Id.* (emphasis in original).

III. DNA DATABASE STATUTES ALLOW DISCLOSURE OF OFFENDERS' DNA

This part will analyze federal and state DNA database statutes to determine whether Congress or the California Legislature contemplated non-government access to forensic DNA profiles. Government officials argue that state and federal law prohibit this access, at least in the context of a criminal trial,²⁴² but a close look at these statutes show that they do not call for total secrecy.

A. THE FEDERAL DNA DATABASE STATUTE DOES NOT REQUIRE TOTAL SECRECY

The FBI built CODIS to link DNA databases across jurisdictions.²⁴³ The FBI learned the importance of interconnectivity from its work with fingerprint databases, which local and state governments maintain in independent and incompatible systems.²⁴⁴ In CODIS, law enforcement officials built a DNA network in which local and state law enforcement agencies could maintain their own records, while still allowing each other to access them.

In 42 U.S.C. § 14132, Congress authorized the FBI director to set up a network of databases containing DNA profiles collected from people convicted of or charged with crimes, or people whose DNA samples are collected under state or local law.²⁴⁵ The statute limits CODIS membership to federal, state, and local criminal justice agencies that keep secret the genetic information they collect.²⁴⁶ However, the statute authorizes disclosure of DNA samples and

²⁴² Ford Motion to Quash, *supra* note 32, at 7, 9; Letter from Dwight E. Adams, Dir., Lab. Div., FBI, to Todd Griffith, Designated State Official, Ariz. Dep't of Pub. Safety (Mar. 17, 2006) (on file with author).

²⁴³ Finn, *supra* note 66.

²⁴⁴ *Id.*

²⁴⁵ 42 U.S.C.A. § 14132(a)(1) (West 2010).

²⁴⁶ 42 U.S.C.A. § 14132(b)(3) (West 2010). The Secretary of Defense is also a member of CODIS. *Id.* Unauthorized disclosure or use of offender DNA information is a crime punishable by a \$250,000 fine or a year in prison. 42 U.S.C. § 14135e(c) (West 2010).

profiles under certain circumstances.²⁴⁷ Generally, those circumstances include: (1) statistical research and quality control, (2) judicial proceedings, (3) criminal trials, and (4) law enforcement use by criminal justice agencies.²⁴⁸ The last category is not an exception that involves disclosure of offenders' DNA and will not be discussed.

The CODIS statute's first confidentiality exception allows disclosure of offenders' DNA for statistical research and quality control purposes, provided that personally identifiable information is removed.²⁴⁹ The statute does not define "quality control," but associates it with "identification research and protocol development" and the creation of "a population statistics database."²⁵⁰ Thus, if the CODIS database really is unsuitable for population genetics research due to duplicates in the system and the presence of family members,²⁵¹ the statute's statistical research and quality control exception authorizes the use of offenders' DNA *samples* to create a new population statistics database from scratch.²⁵² If the offender database turns out to be suitable for research, the statute would seem to authorize such research as a matter of general "quality control" and specific "identification research and protocol development."

The FBI rejects this reading of the quality-control exception. In a letter denying sixteen scientists' request for access to the profiles within CODIS, D. Christian Hassell, director of the FBI laboratory, wrote that this exception is applicable only to criminal justice agencies.²⁵³ For support, Hassell cited a December 2000 resolution of the Federal DNA Advisory Board.²⁵⁴ The Advisory Board said that "samples should only be used by [law enforcement] in accordance

²⁴⁷ 42 U.S.C.A. § 14132(b)(3)(A)–(D) (2010). *See also* 14135e(b) (2010) (referring to "permissive uses" of offender genetic information described in 42 U.S.C. § 14132(b)(3)(A)–(D) (West 2010)).

²⁴⁸ 42 U.S.C.A. § 14132(b)(3)(A)–(D) (West 2010).

²⁴⁹ 42 U.S.C.A. § 14132(b)(3)(D) (West 2010).

²⁵⁰ *Id.*

²⁵¹ *See* Budowle, *supra* note 43, at 60.

²⁵² *See* 42 U.S.C.A. § 14132(b)(3)(D) (West 2010).

²⁵³ Letter from D. Christian Hassell, *supra* note 50, at 1.

²⁵⁴ *Id.*; 42 U.S.C. § 14131(a)(1) (West 2010).

with the DNA Identification Act of 1994.”²⁵⁵ Hassell reads the sentence to say that the samples should be used only by law enforcement, period. But a more natural reading of the sentence shows that “only” modifies “in accordance with the DNA Identification Act of 1994.”²⁵⁶ The Advisory Board was merely emphasizing the primacy of the CODIS statute, not restricting access to CODIS’s contents. Thus law enforcement must follow the CODIS statute, not keep the contents of CODIS to itself.

Hassell also emphasizes that CODIS is limited to the “single use” of law enforcement identification.²⁵⁷ Research into the statistical rarity of DNA profiles conforms with this single use. A Judiciary Committee report that Hassell cites says the CODIS statute “does not purport to resolve” statistical questions.²⁵⁸ Instead, it creates an exception allowing law enforcement and scientists to work together to ensure the integrity of DNA evidence. When the CODIS statute passed, a major problem with DNA evidence was a lack of uniform procedures among laboratories feeding profiles into the system.²⁵⁹ Congress likely wrote the CODIS statute’s quality control exception in general terms in order to give law enforcement the flexibility to respond to future crises of confidence.

The CODIS statute’s second exception “allow[s] disclosure of stored DNA samples and DNA analyses...in judicial proceedings, if otherwise admissible pursuant to applicable statutes or rules.”²⁶⁰ A later statutory section defines “DNA analysis” as the analysis of DNA in a bodily sample, which would seem to refer to DNA profiles, not analyses of databases.²⁶¹ The grammar of the judicial proceedings exception suggests that the limiting “statutes or rules” refer to whether DNA is admissible evidence, not whether other statutes or rules allow DNA to be released in a judicial proceeding at all.²⁶² This

²⁵⁵ Letter from D. Christian Hassell, *supra* note 50, at 2.

²⁵⁶ *See id.*

²⁵⁷ *Id.* at 1–2.

²⁵⁸ *Id.* at 1 (citing H.R. Doc. No. 103-45 (1993)).

²⁵⁹ H.R. Doc. No. 103-45 (1993).

²⁶⁰ 42 U.S.C.A. § 14132(b)(3)(B) (West 2010).

²⁶¹ 42 U.S.C.A. § 14135a(c)(2) (West 2010).

²⁶² *See* 42 U.S.C.A. § 14132(b)(3)(B) (West 2010).

exception puts no limits on who can receive offenders' DNA samples and profiles once they are released in a judicial proceeding. Therefore, this exception likely authorizes the judge-ordered production of offender DNA samples and profiles irrespective of who analyzes them.²⁶³ The only limitation seems to be whether this information is admissible evidence.

The CODIS statute's third exception authorizes some disclosure of forensic DNA to criminal defendants at trial.²⁶⁴ It allows law enforcement to release DNA samples and profiles "for criminal defense purposes, to a defendant, who shall have access to samples and analyses performed in connection with the case in which such a defendant is charged."²⁶⁵ As noted earlier, "DNA analyses" likely refer to DNA profiles.²⁶⁶ Thus while the plain language of the statute reads as though it could include database searches, this subsection likely limits criminal defendants to the DNA samples and profiles that link him or her to the crime charged.²⁶⁷

It could be that this express authorization of disclosure to criminal defendants precludes disclosure of DNA profiles in a criminal case under the judicial proceedings exception. However, Congress would likely have used the words "civil proceedings" rather than "judicial proceedings" if it had intended this result.²⁶⁸ Furthermore, this construction could render the judicial proceedings exception meaningless. Offender DNA profiles are likely relevant to very few judicial proceedings other than criminal trials. They could be of interest to criminal defendants as plaintiffs in civil cases challenging the forensic DNA program's constitutionality, but these suits are discouraged because criminal defendants would have to submit to depositions and lose criminal trial rights. Thus if the CODIS statute limits disclosure of offender DNA profiles in criminal cases to what is allowed under the criminal trials exception, the judicial proceedings

²⁶³ *See id.*

²⁶⁴ 42 U.S.C.A. § 14132(b)(3)(C) (West 2010).

²⁶⁵ *Id.*

²⁶⁶ 42 U.S.C.A. § 14135a(c)(2) (West 2010).

²⁶⁷ *See* 42 U.S.C.A. § 14132(b)(3)(C) (West 2010). This is the FBI's interpretation. Letter from Dwight E. Adams, *supra* note 242 ("The FBI has never interpreted [the criminal defense exception] to permit the defense access to all of the DNA records at the National Index or the performance of any searches.").

²⁶⁸ *See* 42 U.S.C.A. § 14132(b)(3)(B) (West 2010).

language is effectively moot.²⁶⁹ Canons of statutory interpretation argue against this result, leading to the conclusion that the criminal defense exception is the minimum disclosure to which only criminal defendants are entitled, and the judicial proceedings exception is permissive in all cases.²⁷⁰

The FBI rejects this position. In a letter admonishing Arizona's crime lab for complying with a court order in a criminal case authorizing the study the state's offender DNA database, Dwight Adams, director of the FBI's laboratory division, referenced only the criminal defense exception in defining what database information the state can make available to criminal defendants.²⁷¹ While Adams might correctly read the criminal defense exception to bar defendants from running database searches or obtaining copies of the database, his letter ignores the judicial proceedings exception, which authorizes such disclosures if a judge orders them.²⁷² He also ignores the statistical research and quality control exception.²⁷³

Congress expressly authorized population genetics studies of offender databases.²⁷⁴ These studies could be done in a variety of ways, according to the CODIS statute—through internal quality control, outside research, or within judicial proceedings.²⁷⁵ This article advocates the outside research approach in order to ensure credible results and insulate researchers from the pressures of the adversarial criminal justice system.²⁷⁶ However, the federal CODIS

²⁶⁹ See 42 U.S.C.A. § 14132(b)(3)(B)–(C) (West 2010).

²⁷⁰ See 42 U.S.C.A. § 14132(b)(3)(C) (West 2010).

²⁷¹ Letter from Dwight E. Adams, *supra* note 242.

²⁷² *Id.*

²⁷³ *Id.*

²⁷⁴ See 42 U.S.C.A. § 14132(b)(3)(D) (West 2010).

²⁷⁵ See 42 U.S.C.A. § 14132(b)(3)(B), (D) (West 2010).

²⁷⁶ The NAS acknowledged these pressures by recommending the creation of a new agency outside of the Department of Justice (DOJ) to control forensic science funding and research. NAS REPORT, *supra* note 54, at 19. Professor Paul Giannelli calls this a “controversial but needed reform.” Paul C. Giannelli, Daubert and Forensic Science: The Pitfalls of Law Enforcement Control of Scientific Research 6 (Case Research Paper Series in Legal Studies, Paper No. 2010-6, 2010). DOJ, the FBI Crime Lab, and some prosecutors “have attempted to shape science by controlling the research agenda, hiding unwelcome test results, attacking legitimate studies that were considered unfavorable, harassing

statute allows all judge-supervised database research as long as the results are admissible evidence.²⁷⁷

B. CALIFORNIA'S DNA DATABASE LAW DOES NOT REQUIRE TOTAL SECRECY

Notwithstanding possible avenues of disclosure under federal law, state law governs state databases. California law on this point is similar to federal law: it commands that all DNA profiles and samples be kept secret unless certain exceptions apply.²⁷⁸ These exceptions include: (1) administrative or judicial proceedings;²⁷⁹ (2) training, research, statistical analysis of populations, quality assurance, or quality control;²⁸⁰ (3) criminal trials;²⁸¹ and (4) law enforcement use.²⁸² The last category does not involve disclosure of DNA evidence outside of law enforcement, and will not be discussed here.

Like the federal statute, California law allows disclosure of DNA database information in judicial proceedings.²⁸³ However, the California statute leaves such disclosures to the discretion of law enforcement as long as the information disclosed “pertains to the basis” for identifying, arresting, investigating, prosecuting, or excluding a person related to the case.²⁸⁴ The exception explicitly forbids courts from compelling disclosure of “any” database information “in a criminal or civil proceeding ... whether by subpoena

scientists who disagreed with them, and ‘spinning’ these issues in the press.” *Id.* at 7. Professor Giannelli’s article chronicles these misdeeds.

²⁷⁷ See 42 U.S.C.A. § 14132(b)(3)(B) (West 2010).

²⁷⁸ See CAL. PENAL CODE § 299.5(a) (West 2010). Unauthorized use or disclosure of DNA profiles is a crime punishable by fine or imprisonment. CAL. PENAL CODE § 299.5(i) (West 2010).

²⁷⁹ CAL. PENAL CODE § 299.5(k) (West 2010); CAL. PENAL CODE § 299.5(h) (West 2010).

²⁸⁰ CAL. PENAL CODE § 299.5(m) (West 2010). See also CAL. PENAL CODE § 299.6(a)(5) (West 2010).

²⁸¹ CAL. PENAL CODE § 299.5(g), (j) (West 2010).

²⁸² CAL. PENAL CODE § 299.5(f) (West 2010).

²⁸³ CAL. PENAL CODE § 299.5(k) (West 2010).

²⁸⁴ *Id.*

or discovery.”²⁸⁵ This likely blocks judge-ordered searches or supervised defense-expert research, while leaving open such measures DOJ should undertake voluntarily.²⁸⁶ Yet voluntary action seems unlikely given DOJ’s opposition to DNA database research.²⁸⁷

The California statute also mirrors federal law by allowing disclosure of offender database information for research and quality control purposes.²⁸⁸ California’s DNA database law allows DOJ to use “anonymous records or criminal history information” for training, research, statistical analysis of populations, quality assurance, or quality control.²⁸⁹ The law says that DOJ, its agent, or a local public laboratory can use this information for statistical analysis, authorizing DOJ to conduct the research itself, or to allow outsiders to do it.²⁹⁰ If it turns out that DNA profiles and samples do not qualify as “records or criminal history information,” a later statutory section makes clear that DOJ has the discretion to share or disseminate a range of information relating to offender DNA databases, including DNA profiles.²⁹¹ Under this section, state authorities could deliver the offender database to “any third party that [DOJ] deems necessary to assist the department’s crime laboratory with statistical analysis of population databases.”²⁹² Therefore, California law allows DOJ to conduct statistical research of the database, whether internally or by an outside researcher.

The third exception governs criminal trials.²⁹³ State authorities are required to release DNA and other forensic identification information to a defendant in preparation for trial.²⁹⁴ Unlike the

²⁸⁵ CAL. PENAL CODE § 299.5(h) (West 2010) (noting that the purpose of secrecy is to “protect the confidentiality and privacy of database and data bank information”).

²⁸⁶ *See id.*

²⁸⁷ *See e.g.*, Ford Motion to Quash, *supra* note 32; Davis Motion to Quash, *supra* note 158.

²⁸⁸ CAL. PENAL CODE § 299.5(m) (West 2010); CAL. PENAL CODE § 299.6(a)(5) (West 2010).

²⁸⁹ CAL. PENAL CODE § 299.5(m) (West 2010).

²⁹⁰ *See id.*

²⁹¹ CAL. PENAL CODE § 299.6(a) (West 2010).

²⁹² CAL. PENAL CODE § 299.6(a)(5) (West 2010).

²⁹³ CAL. PENAL CODE § 299.5(g), (j) (West 2010).

²⁹⁴ CAL. PENAL CODE § 299.5(g), (j) (West 2010).

federal statute, where the express authorization of disclosure to criminal defendants in a criminal trial could preclude their use of the judicial proceedings exception, California's judicial proceedings exception mentions the "grand jury" and "prosecution," making clear that it applies to criminal trials.²⁹⁵ Thus California's criminal trials exception, which commands law enforcement to release some DNA information, is the baseline of disclosure upon which the judicial proceedings exception builds.

If DOJ does not use its discretion to allow scientists to study the rarity of DNA profiles using California's offender database, the constitutionality of California's DNA database statute could be questioned.²⁹⁶ Unlike the federal statute, which allows courts to release DNA database information in judicial proceedings,²⁹⁷ California law leaves such disclosures to the discretion of law enforcement.²⁹⁸ It thereby blocks the only external means of addressing the deficiencies of (1) using databases to identify defendants without providing them a means to refute the identification; and (2) presenting rarity statistics that might not be accurate. If DOJ refuses to use its discretion to allow third-party evaluations of DNA databases, the portion of California's database law that prohibits court-ordered database research could be struck down.²⁹⁹

IV. AGENCIES SHOULD RELEASE OFFENDER DNA PROFILES TO RESEARCHERS USING SAFEGUARDS IN PLACE IN OTHER CONTEXTS

As Part III demonstrates, the dearth of statistical research using offenders' DNA profiles is a matter of law enforcement prerogative rather than law. This part explores three contexts in which government discloses sensitive information about individuals: (1) through the California Public Records Act, (2) during trials, and (3) to independent researchers. It then evaluates each and determines that

²⁹⁵ See CAL. PENAL CODE § 299.5(k).

²⁹⁶ See *supra* note 48.

²⁹⁷ See 42 U.S.C.A. § 14132(b)(3)(B) (West 2010).

²⁹⁸ See CAL. PENAL CODE § 299.5(k) (West 2010).

²⁹⁹ See CAL. PENAL CODE § 299.5(m) (West 2010); CAL. PENAL CODE § 299.6(a) (West 2010).

the independent research model is the most suitable for scientific study of offenders' DNA profiles.

A. THE CALIFORNIA PUBLIC RECORDS ACT COMPELS DISCLOSURE OF ARIZONA SEARCH RESULTS

The California Public Records Act can be a powerful tool for discovering information generally considered private. For example, it has been used to compel disclosure of state and local government payroll information, identifying public employees with the money they make.³⁰⁰ It can penetrate personnel files, unveiling misconduct complaints against public employees.³⁰¹

However, the California Public Records Act is not boundless. While it "declares that access to information concerning the conduct of the people's business is a fundamental and necessary right," it exempts some information from disclosure.³⁰² Law enforcement intelligence and investigatory information is one exemption.³⁰³

The California DOJ cited the law enforcement disclosure exemption when it rejected a Public Records Act request the author submitted seeking information about California's offender DNA database.³⁰⁴ The author asked the state for "records and documentation" of DNA matches between individuals within the offender database; these would include numerical and profile results of an Arizona search, a chance discovery in the course of an

³⁰⁰ See, e.g., Press Release, UCSF, UC Employee Compensation Made Available to Public (Oct. 10, 2007), <http://today.ucsf.edu/stories/uc-employee-compensation-made-available-to-public/>.

³⁰¹ See, e.g., *Bakersfield City Sch. Dist. v. Superior Court*, 118 Cal. App. 4th 1041 (Cal. Ct. App. 2004) (upholding a lower court decision to make public sexual misconduct complaints against a former middle school vice principal suspected of killing his family).

³⁰² CAL. GOV'T CODE § 6250 (West 2010).

³⁰³ CAL. GOV'T CODE § 6254(f) (West 2010).

³⁰⁴ California courts have not yet ruled on the specific question of whether offenders' genetic information—samples or profiles or both—is subject to disclosure under the Act. However, the Supreme Court of Virginia held that DNA samples from a rape case were not subject to disclosure under the Virginia Freedom of Information Act. *Globe Newspaper Co. v. Com.*, 570 S.E.2d 809 (Va. 2002). In that case, several newspapers seeking to re-test DNA from a decades-old rape and murder for which a man was convicted and put to death petitioned the court for access to the evidence. *Id.* at 810–11.

investigation, or quality control research.³⁰⁵ If these records “were to exist,” DOJ said, they would be exempt from disclosure under the Public Records Act.³⁰⁶ What follows is an analysis of DOJ’s arguments for rejecting the author’s request.

First, DOJ said DNA match information is “confidential and exempt from disclosure.”³⁰⁷ In *ACLU Foundation v. Deukmejian*, the California Supreme Court held that law enforcement data consisting of “personal identifiers, confidential sources, and confidential information relating to criminal activity” is exempt intelligence information under the California Public Records Act.³⁰⁸ There the ACLU was concerned about an index card system police used to monitor organized crime, a system that tracked suspects’ “associates,” such as family members, business acquaintances, and their attorneys.³⁰⁹ Each card listed the associate’s name, alias, occupation, family members, vehicles, associates, arrests, modus operandi, and physical traits.³¹⁰ This information was exempt from disclosure under the California Public Records Act, the court held, as is (1) any “information which might lead the knowledgeable or inquisitive to infer the identity of the individual in question,” or (2) “information supplied in confidence.”³¹¹ The court based its decision in part on the fact that while the ACLU sought to uncover government abuses, the records could also be used for “less noble purposes.”³¹²

Under *ACLU Foundation*, DOJ is likely correct that the California Public Records Act prohibits general disclosure of offender DNA

³⁰⁵ Letter from Jill Spriggs, Chief, Bureau of Forensic Servs., Cal. Dep’t of Justice, to author (July 28, 2008) (on file with author). The author did not ask for a copy of the contents of California’s entire offender database. *Id.* Given the state’s denial of the author’s request for matching profiles, or numbers of matches, this request would likely have been denied.

³⁰⁶ *Id.*

³⁰⁷ *Id.*

³⁰⁸ *ACLU Found. v. Deukmejian*, 32 Cal.3d 440, 452 (Cal. 1982).

³⁰⁹ *Id.* at 444 (noting that in attacking the index card system, the ACLU “cites a striking example of the potential for abuse in unmonitored gathering of information by law enforcement agencies”). For example, a state senator was listed as an associate of a Black Panther Party member who had rented the senator’s house for four months. *Id.*

³¹⁰ *Id.*

³¹¹ *Id.* at 450.

³¹² *Id.* at 451.

profiles. First, offenders supply them to law enforcement in confidence insofar as DNA collection statutes limit their disclosure to circumstances that do not include broad public dissemination. Second, they identify people, albeit to a lesser degree than the name or alias records in *ACLU Foundation*.³¹³ And law enforcement officials have long argued that DNA profiles are junk, useful for identification only if names are attached.³¹⁴ However, even if offenders' profiles were stripped of identifying information, someone with the right equipment and a copy of the offender database could swab a coffee mug, profile the drinker's DNA, and search for a match. In California, the most this would reveal is whether police had arrested the coffee drinker for a felony. Despite the fact that adult arrest records are public,³¹⁵ this is a revelation likely prohibited by the court's construction of the California Public Records Act.

DOJ also objected to revealing whether the offender DNA database contained matching DNA profiles.³¹⁶ This information is not the kind of information kept secret in *ACLU Foundation* because there is nothing personally identifying about the fact of a match.³¹⁷ Furthermore, while DNA samples are provided to the government in limited confidence, the number of matching profiles within the offender database is not supplied to government—it is discovered *by* government.

DOJ's second objection to the request was that a record of a DNA match is exempt investigatory information because it is "a potential lead in a case unsolved by other means."³¹⁸ While DNA profiles themselves are potential leads, the fact that the database contains matches between individuals is not. Just because the DNA database is fundamentally an investigatory tool, matches or associations within it

³¹³ See *id.* at 444.

³¹⁴ See *supra* notes 198–205 and surrounding text.

³¹⁵ The Los Angeles Police Department website lists the various kinds of information available to the public under the California Public Records Act. LAPD, California Public Records Act, http://www.lapdonline.org/home/content_basic_view/36329 (last visited Apr. 17, 2010). The site lists arrest logs, 911 recordings, crime reports, accident reports, sex offender registry files, local criminal history for an individual, warrant information, and some juvenile records. *Id.*

³¹⁶ See Letter from Jill Spriggs, *supra* note 305, at 2.

³¹⁷ See *ACLU Found.*, 32 Cal. 3d at 444.

³¹⁸ Letter from Jill Spriggs, *supra* note 305.

are not necessarily records of an investigation.³¹⁹ If DOJ learned of database matches while conducting an investigation, these matches could be confidential records of that investigation, but if it learned of them while conducting an Arizona search or statistical studies, they would not be.

Finally, DOJ said that DNA database records are exempt from disclosure under the Public Records Act provision that protects information otherwise kept secret by federal or state statute.³²⁰ DOJ says federal and state laws prohibit disclosure of DNA profiles and matches within an offender database.³²¹ While both levels of government prohibit disclosure of DNA profiles and samples to the public, precluding such a disclosure under the California Public Records Act, the fact that matches exist *at all* is not secret.³²² If such matches exist, the fact that they do is a matter of public record under California law.

B. SOME FINGERPRINT AND DNA SEARCH RECORDS ARE RELEASED TO DEFENDANTS

Fingerprinting was the premier criminal identification procedure before DNA came along.³²³ Fingerprints, like DNA profiles, are stored in state-run databases, and technicians search them using the Automated Fingerprint Identification System (AFIS).³²⁴ A fingerprint examiner scans latent prints lifted from crime scenes and uses computer software to mark identifying characteristics of the print—where ridges change direction, whorl, or loop.³²⁵ The positioning of these markings is what the software uses to hunt through the database, searching through prints of known offenders that have a

³¹⁹ *Id.* at 2–3.

³²⁰ *Id.* at 3 (citing CAL. GOV'T CODE § 6254(k) (West 2010)).

³²¹ *Id.*

³²² *See id.*

³²³ *See, e.g.,* Jennifer L. Mnookin, *Fingerprint Evidence in an Age of DNA Profiling*, 67 BROOK. L. REV. 13, 39–43 (2001).

³²⁴ U.S. DEP'T OF JUSTICE, OFFICE OF THE INSPECTOR GEN., A REVIEW OF THE FBI'S HANDLING OF THE BRANDON MAYFIELD CASE, 118 (Mar. 2006), available at http://www.usdoj.gov/oig/special/so601/PDF_list.htm.

³²⁵ *Id.* at 119.

similar constellation of ridge markings.³²⁶ AFIS produces a list of possible matches, and the lab technician sorts through them, eliminating fingerprints that have a whorl where the crime-scene print has a loop.³²⁷ If an AFIS-produced print survives this round of examination, the lab technician pulls original print cards and begins a ridge-by-ridge comparison between the known print and the latent print lifted from the crime scene.³²⁸

Criminal defense attorneys representing clients identified through AFIS have argued that the list of prints flagged through an AFIS search is exculpatory and thus discoverable under *Brady v. Maryland*.³²⁹ Courts have rejected this argument,³³⁰ but a trial court in New Jersey held that the results of an AFIS analysis qualify as a scientific test under the state's criminal discovery rules, and must therefore be turned over to the defense.³³¹ The defendant's goal in seeking AFIS results is to show "that there are other persons with a match that is as close" to the crime-scene print as that of the defendant.³³² The court deemed this "relevant evidence [that] must be given to the defendant."³³³

³²⁶ *Id.*

³²⁷ *Id.* at 120.

³²⁸ *Id.*

³²⁹ *Brady v. Maryland*, 373 U.S. 83, 87 (1963); See Forensics Library, National Legal Aid & Defender Association, http://www.nlada.org/Defender/forensics/for_lib/Index/Fingerprints/showHelp (last visited Apr. 16, 2010) (recommending that defense attorneys file discovery requests for AFIS results. Relevant questions to ask include: "Did the AFIS place your client's print in the first position? If not, have the examiner explain why not? [sic] Does the examiner know better than the computer? Have your defense examiner look at the database print that did appear in the first position. Was the AFIS search run more than once? Were there different results? What explains the different results?").

³³⁰ See, e.g., *State v. Green*, No. 90092, 2008 WL 2623959, at *2 (Ohio Ct. App. July 3, 2008); *People v. Tims*, No. A092799, 2002 WL 202477, at *11 (Cal. Ct. App. Feb. 7, 2002).

³³¹ *State v. Feldman*, 604 A.2d 242, 244 (N.J. Super. Ct. Law Div. 1992). Despite the importance of this evidence and the fact that it "must be given to the defendant," the court held that its destruction was not in bad faith and that its unavailability was neither material to nor did it prejudice the defense. *Id.* at 245.

³³² *Id.*

³³³ *Id.*

Turning to DNA, some courts have ordered Arizona searches, but most have stopped short of releasing to defense attorneys all profiles contained in offender DNA databases.³³⁴ However, in 1999, a trial court judge in Colorado ordered the state to deliver the contents, without names, of its 5,629-person database to defense attorneys as discovery.³³⁵ The Colorado database was too small to do the kind of large-scale statistical testing urged by scientists.

In California, attorneys defending DNA database cases can request their clients' "hit file," which documents the DNA database search that led authorities to the defendant.³³⁶ The hit file contains information about how the DOJ lab handled the defendant's DNA, when it uploaded the profile to the database system, when the local lab uploaded the crime-scene sample, the letter DOJ sent to the local crime lab notifying it of a hit, and records of DOJ's re-analysis of the defendant's sample to double-check the match.³³⁷ In rare cases, a database search identifies more than one offender as a potential match to the crime-scene profile, said DOJ attorney Michael Chamberlain.³³⁸ Like AFIS, DNA database software does not look only for perfect matches.³³⁹ In the case of more than one matching profile, defense attorneys would find out through the hit file, according to Chamberlain.³⁴⁰ Additionally, he observed that if a defense attorney questioned the local lab's determination that other offenders identified through the database search could not have contributed the crime-scene sample, the attorney could seek these offenders' profiles in discovery.³⁴¹ Whether profiles determined to be non-matches are discoverable is adjudicated on a case-by-case basis. The disclosure of profiles would allow defense attorneys to see for themselves why state DNA analysts determined that the defendant's profile matched the

³³⁴ See *supra* note 32.

³³⁵ E-mail from Steve Jacobson, *supra* note 32.

³³⁶ E-mail from Michael Chamberlain, Deputy Attorney Gen., Cal. Dep't of Justice, to author (Mar. 11, 2010 10:40 PST) (on file with author).

³³⁷ *Id.*

³³⁸ *Id.*

³³⁹ *Id.*

³⁴⁰ *Id.*

³⁴¹ *Id.*

crime-scene evidence and the others' did not. Finally, prosecutors also release DNA profiles, with names, where police—not databases—identify multiple suspects, test their DNA, and find that only one matches the crime-scene profile.³⁴²

The disclosures mentioned above, while potentially useful to individual defendants, do not reach the integrity of DNA statistics. The next part examines how California uses sensitive data sets to conduct large-scale research.

C. GOVERNMENT AGENCIES DISSEMINATE INFANTS' DNA TO MEDICAL RESEARCHERS

Government agencies share sensitive information about individuals in the context of medical and social science research.³⁴³ Again, this article will use California as a case study.

In California, government agencies that collect personal information may release it to the University of California or other non-profit educational institutions for research purposes.³⁴⁴ These releases must be approved by the Committee for the Protection of Human Subjects (CPHS), a division of the California Health and Human Services Agency.³⁴⁵ A snapshot of CPHS-approved studies ongoing as of June 2008 shows CPHS overseeing 449 projects in which researchers were using state databases containing personal information about individuals.³⁴⁶ Of about two dozen genetic research projects among them, nine used state-collected blood samples from newborns.³⁴⁷ The newborn screening program the only state-run

³⁴² E-mail from Michael Chamberlain, Deputy Attorney Gen., Cal. Dep't of Justice, to author (Mar. 12, 2010 12:01 PST) (on file with author).

³⁴³ See E-mail from Roxana Killian, Adm'r, Comm. for the Protection of Hum. Subjects (CPHS), to author (Feb. 25, 2009 14:31 PST) (on file with author) (including attached spreadsheets showing 449 research projects using state databases as of June 2008; nearly all of them are medical or social science studies).

³⁴⁴ CAL. CIVIL CODE § 1798.24(t) (West 2010).

³⁴⁵ *Id.*

³⁴⁶ See E-mail from Roxana Killian, *supra* note 380.

³⁴⁷ Lisa Croen, Prenatal and Neonatal Biologic Markers for Autism (July 2, 2008) (unpublished research protocol for study by Kaiser Permanente Div. of Research) (on file with author); Judith K. Grether, Candidate Genes for Childhood Autism (July 13, 2007) (unpublished research protocol for study by Cal. Dep't of Public Health) (on file with author); Martin Kharrazi, Molecular Markers of Preterm Birth—Maternal and Infant

genetic registry in which participation is mandatory; state law says all parents must allow their children to give a blood sample unless they object on religious grounds.³⁴⁸ Once collected, the samples belong to the state, which may release them to state agencies and academic or nonprofit researchers seeking to do certain kinds of health research.³⁴⁹ If researchers breach confidentiality requirements, the person affected can seek as much as \$10,000 in compensatory and civil damages, plus attorney fees and the cost of litigation.³⁵⁰

Researchers with a department's permission to analyze individuals' personal information, including genetic samples from the newborn screening program, must also seek approval and subsequent review from CPHS.³⁵¹ Organizations eligible for approval are limited

Genetic Contributions to Preterm Birth: The Inflammatory Response (Aug. 25, 2008) (unpublished research protocol for study by Cal. Dep't of Public Health) (on file with author); Martin Kharrazi, An Examination of Unexplained Courses of Early Child Mortality and the Value of Newborn Screening (July 22, 2008) (unpublished research protocol for study by Cal. Dep't of Public Health) (on file with author); Martin Kharrazi, An Assessment of the Incidence of Cytomegalovirus Infection in Newborns in California (Feb. 28, 2008) (unpublished research protocol for study by Cal. Dep't of Public Health) (on file with author); Martin Kharrazi, Tobacco Exposure in Pregnant Women in Minority Populations (Mar. 2, 2004) (unpublished research protocol for study by Cal. Dep't of Public Health) (on file with author); Martin Kharrazi, An Assessment of the Seroprevalence of HCV Infection in Childbearing Women in San Diego, Orange and Imperial Counties (Sept. 4, 2002) (unpublished research protocol for study by Cal. Dep't of Public Health) (on file with author); Roberta McKean-Cowdin, Gene-Environment Factors in Childhood Brain Tumors (Jan. 2007) (unpublished research protocol for study by Univ. of S. Cal. Norris Cancer Ctr.) (on file with author); Gayle Windham, Autism and Peri-natal Hormone Markers (July 6, 2006) (unpublished research protocol for study by Cal. Dep't of Public Health) (on file with author).

³⁴⁸ CAL. HEALTH & SAFETY CODE §§ 124975(j), 125000 (West 2010); Bradford L. Therrell et al., *Status of Newborn Screening Programs in the United States*, 117 PEDIATRICS S212, S230 (2006). All states and the District of Columbia collect newborns' blood to screen for diseases, and each state imposes its own consent and opt-out requirements. *Id.* at S212.

³⁴⁹ *Id.* at S231.

³⁵⁰ *Id.*

³⁵¹ Until 2005, CPHS oversaw only Health and Human Services Agency research. See *Governor Schwarzenegger Signs Legislation to Fight Identity Theft*, STATES NEWS SERVICE, Sept. 22, 2005. In 2005, the California legislature amended the Information Practices Act to require CPHS to approve scientific research proposals before any agency discloses personal information to researchers. *Id.* The impetus for this change came in 2004, when a visiting economics researcher at UC Berkeley was hacked, potentially compromising 600,000 Californians' Social Security numbers and other personal data. Keay Davidson, *BERKELEY: Researcher Says Hacked Data Stored Properly*, S.F. CHRON., Oct. 23, 2004, at B2. The researcher was studying the impact of wages and benefits upon the state's ability to recruit and retain in-home health care workers. *Id.*

to the University of California or a nonprofit educational institution conducting scientific research.³⁵² To get CPHS's approval, the researcher must: (1) submit a plan that includes administrative, physical, and technical safeguards designed to protect personal data from improper use and disclosure;³⁵³ (2) submit a plan to destroy or return all personal information when it is no longer needed, or a long-term plan for protecting the information;³⁵⁴ and (3) promise in writing that the personal information will not be reused, disclosed, or used in an unauthorized manner.³⁵⁵ CPHS then determines whether the information requested is the minimum necessary to do the research.³⁵⁶ If the research requires Social Security numbers, the law says they must be encrypted.³⁵⁷ Moreover, if feasible—meaning if cost, time, and technical expertise permit—California law requires the state agency holding the data to process it for researchers in order to minimize the disclosure of personal information.³⁵⁸

Unlike the public records and criminal trial contexts, California's system of disseminating newborn genetic information depends upon the cooperation of government agencies. It does not contain a mechanism through which state officials might be compelled to consider whether offender DNA profiles are analogous to infants' DNA samples and should be released in the same way. However, a court could decide that the constitutional rights of defendants require large-scale statistical research of offender DNA databases, or law enforcement agencies could follow Australia's lead by releasing offender databases to outside researchers for study. The next section examines how best to handle DNA profiles if they are released.

³⁵² CAL. CIVIL CODE § 1798.24(t)(1) (West 2010).

³⁵³ CAL. CIVIL CODE § 1798.24(t)(1)(A) (West 2010).

³⁵⁴ CAL. CIVIL CODE § 1798.24(t)(1)(B) (West 2010).

³⁵⁵ CAL. CIVIL CODE § 1798.24(t)(1)(C) (West 2010).

³⁵⁶ CAL. CIVIL CODE § 1798.24(t)(2)(A)–(C) (West 2010).

³⁵⁷ CAL. CIVIL CODE § 1798.24(t)(2)(D) (West 2010).

³⁵⁸ CAL. CIVIL CODE § 1798.24(t)(2)(E) (West 2010).

D. NEWBORN GENETIC RESEARCH PROVIDES THE BEST MODEL FOR
RELEASING SENSITIVE INFORMATION, AND COURTS OR LAWMAKERS
SHOULD IMPLEMENT A SIMILAR PROCEDURE FOR DNA

The most appropriate way for experts to study offender DNA profiles is to release them through a process similar to California's system of facilitating medical and social science research. While some information about offender databases should be made available through open-records laws, DNA profiles are personally identifying enough to warn against entrusting them to the general public. Furthermore, access to the statistical information contained in criminal offender databases should not depend on the insistence of a defendant's lawyer, the quality of a defense expert, or the whim of a single judge.

Criminal trials are the most convenient places to raise questions about the rarity of DNA profiles,³⁵⁹ but they are not designed to answer them. The history of forensic DNA evidence is marked by mistakes, and "for nearly every laboratory mistake or malfeasant act, there were lawyers and judges who failed to catch it."³⁶⁰ As Law Professor Erin Murphy points out, the technical complexity of DNA evidence makes "close and continuous judicial scrutiny of their methodological soundness less likely."³⁶¹ Well-meaning judges might be intimidated by the science involved, or wary of taking action that is seemingly contrary to the vast majority of courts that have accepted forensic DNA as legitimate courtroom evidence.³⁶² In addition,

³⁵⁹ See *supra* note 32.

³⁶⁰ Murphy, *supra* note 37, at 767. Professor Erin Murphy points out courtroom missteps, including an expert testifying that he had never heard of a case in which unrelated individuals matched at three or four DNA markers. *Id.* She also notes that in Virginia and Texas, people were "wrongly jailed ... for years on the basis of falsely inculcating DNA evidence." *Id.* (citing, for example, Steve McVicker, *More DPS Labs Flawed: DNA Testing Woes Across State Threaten Thousands of Cases*, HOUSTON CHRON., Mar. 27, 2004, at A1; Adam Liptak, *You Think DNA Evidence is Foolproof? Try Again*, N.Y. TIMES, Mar. 16, 2003, §4, at 5).

³⁶¹ *Id.* at 768.

³⁶² *Id.* at 769. In *United States v. Plaza*, withdrawn from bound volume but available at 2002 WL 27305 (E.D. Pa. 2002), a trial court held some fingerprint evidence to be insufficiently reliable. *Id.* The trial court encountered so much ridicule that it reversed itself despite significant scholarship supporting its initial position. *Id.* at 768 n.205 (citing *United States v. Plaza*, 188 F. Supp. 2d 549, 576 (E.D. Pa. 2002) (vacating earlier opinion)).

releasing the offender database to individual criminal defendants, while potentially useful to their defense,³⁶³ is inefficient because it burdens courts and attorneys without taking on what is potentially a systemic problem: whether DNA statistics used in all DNA cases are correct.

Bicka Barlow, a San Francisco defense attorney who initiated the original Arizona search litigation and continues to push for access to California's offender profiles, acknowledges that lawsuits are not ideal for answering the questions raised by DNA evidence.³⁶⁴ Battling it out in court "should not necessarily be how it works," she said.³⁶⁵ The inertia of the courts make them slow to confront possible problems with DNA evidence, she said, and "so much of [whether the issue gets raised and whether judges take it seriously] depends on the quality of the lawyering."³⁶⁶

If statistical research of offender DNA profiles is left to the adversarial criminal justice system, answers to questions about the rarity of DNA evidence will be slow in coming, if at all. Courts might not feel comfortable questioning DNA evidence, and defense attorneys seldom have the expertise to challenge it. Even if attorneys did have the expertise, they might "reasonably conclude that it requires too great an effort, and reaps too little a reward, to study such evidence in the hopes of uncovering a flawed methodological approach."³⁶⁷ As Professor Erin Murphy points out, defense attorneys' time might be better spent working on a defense compatible with the evidence, or negotiating a good plea deal for their client.³⁶⁸

If an individual criminal case is not the appropriate place to deal with questions about DNA evidence, universities are. Medical researchers interested in the ethical issues involved in conducting

³⁶³ For example, defense experts could search the database for near misses and argue that a relative of the near miss is plausibly the perpetrator. Or the defense could search the database for other matches in an attempt to show that DNA matches are more common than DNA statistics seem to suggest.

³⁶⁴ Telephone Interview with Bicka Barlow, Attorney, S.F. Office of the Pub. Defender (Feb. 24, 2009). Barlow was one of John Puckett's defense attorneys. Smith, *supra* note 4, at 66.

³⁶⁵ Telephone Interview with Bicka Barlow, *supra* note 364.

³⁶⁶ *Id.*

³⁶⁷ Murphy, *supra* note 37, at 770.

³⁶⁸ *Id.* at 771.

genetic research on newborn DNA samples acknowledge that “concern for confidentiality continues to fuel the ethical debate over public health benefits versus personal privacy.”³⁶⁹ Substitute “criminal justice benefits” for “public health benefits,” and this statement encapsulates the issue at hand: whether assuring the integrity of the criminal justice system outweighs possible privacy issues involved in allowing outside researchers to study offender DNA databases. It does. In California, the public health benefits of releasing newborns’ DNA samples to researchers outweigh infants’ personal privacy.³⁷⁰ Given that courts have consistently held that offenders have a diminished privacy interest in their identifying information,³⁷¹ the countervailing public interest in the accuracy of DNA statistics wins out here.³⁷²

It is true that neonatal blood samples are collected at least in part for their research potential, while offenders’ DNA samples are not. Professors David Lazer and Viktor Mayer-Schonberger argue against using offenders’ DNA in ways not intended at the time of collection.³⁷³ However, some collateral uses, such as quality control, “certainly support the core purpose of the collected data,” they say.³⁷⁴ Such uses call for “flexible rules...coupled with oversight mechanisms.”³⁷⁵

³⁶⁹ Bradford L. Therrell et al., *Guidelines for the Retention, Storage and Use of Residual Dried Blood Spot Samples After Newborn Screening Analysis: Statement of the Council of Regional Networks for Genetic Services*, 57 *BIOCHEMICAL & MOLECULAR MED.* 116, 120 (1996).

³⁷⁰ See Therrell et al., *supra* note 348, at S231.

³⁷¹ See, e.g., *United States v. Kreisel*, 508 F.3d 941, 948 (9th Cir. 2007); *United States v. Weikert*, 504 F.3d 1, 12–13 (1st Cir. 2007); *United States v. Kincade*, 379 F.3d 813, 835, 837–38 (9th Cir. 2004).

³⁷² Professor David Kaye argues that while the right to informed consent is implicated in the collection of personal information for medical or scientific research, this right “does not apply when samples are legally compelled and the information extracted from them is used solely to ensure that the very system that justifies this disclosure is working as it should.” Kaye, *supra* note 33, at 30.

³⁷³ See David Lazer & Viktor Mayer-Schonberger, *Statutory Frameworks for Regulating Information Flows: Drawing Lessons for the DNA Data Banks from Other Government Data Systems*, 34 *J.L., MED. & ETHICS* 366, 372 (2006).

³⁷⁴ *Id.*

³⁷⁵ *Id.*

Flexible rules and oversight mechanisms characterize California's system for allowing scientists to study newborns' DNA samples.³⁷⁶ The system is flexible because CPHS approves research protocols on a case-by-case basis, evaluating each project's proposed means of protecting sensitive information about individuals. Oversight is achieved through regular reports to CPHS. A review of these reports shows that they address the same issues that tend to stymie research of offender databases—privacy and security.³⁷⁷ For example, University of Southern California scientists using genetic information to study the environmental factors associated with childhood brain tumors assured CPHS in their report that while loss of confidentiality is a possibility, the scientists use intensive security measures.³⁷⁸ These include passwords, locked cabinets, allowing access to information on a need-to-know basis, and a practice of storing genetic information separately from identifying information.³⁷⁹ Even if a breach were to occur, the researchers wrote, their work examines portions of the genome that do not code for stigmatizing traits.³⁸⁰

If these kinds of measures are sufficiently protective of genetic samples, they should be more than adequate to safeguard the release of DNA profiles—strings of numbers useful for little more than differentiating among individuals.³⁸¹ In other words, if CPHS oversight is not enough to protect offender profiles, it is not enough to protect children's DNA samples. If that is the case, all genetic research using newborns' samples should stop until the state has enough confidence in its means of protecting privacy to allow genetic research of every kind to continue. This would be devastating to important genetic research, but the state cannot tout its privacy system with regard to newborns and dismiss it as inadequate for dealing with known offenders.

³⁷⁶ While this article assumes CPHS oversight to be as robust as it appears on paper, it is plagued by low staffing levels. See E-mail from Roxana Killian, Adm'r, CPHS (Feb. 27, 2009, 09:34 PST) (on file with author).

³⁷⁷ See *supra* Part II(B).

³⁷⁸ Roberta McKean-Cowdin, Gene-Environment Factors in Childhood Brain Tumors 3 (Jan. 2007) (unpublished research protocol, on file with author).

³⁷⁹ *Id.*

³⁸⁰ *Id.*

³⁸¹ Even if DNA profiles turn out to code for traits, the most an offender database will reveal is whether offenders do or do not have those traits.

The difference between the way California handles newborns' genetic information and offenders' DNA profiles is likely its relative desire to see each type of research done. This is a common phenomenon where government agencies hold data on individuals, said Franklin E. Zimring, professor at UC Berkeley School of Law. "Researchers have no trouble [getting information] when [agencies] want to give it," he said.³⁸² But when the agency does not want to release the information, outsiders have very few avenues for compelling disclosure, he said.³⁸³ This is especially true in cases in which privacy rights are at stake.³⁸⁴ "If you create a privacy right for individuals, you are essentially giving a club to administrative agencies to simply refuse," he said.³⁸⁵

An agency's decision to release information "can go either way," Zimring said, and when he sought to study the impact of California's 1994 "Three Strikes and You're Out" legislation, the decision went his way.³⁸⁶ The Three Strikes law greatly enhanced punishments for recidivists, and the Public Policy Institute of California awarded Zimring a grant to study whether these punishments deterred repeat offenders.³⁸⁷ To do the study, Zimring needed a list of arrestees and their alleged crimes before Three Strikes, and a list of them after. In response to his request, the state DOJ gave Zimring 3000 criminal records, including names and other identifying information.³⁸⁸ Zimring says he got the information because people at state agencies "thought it was interesting research."³⁸⁹ He is not always so lucky: "If

³⁸² Interview with Franklin E. Zimring, Professor, Univ. of Cal., Berkeley Sch. of Law, in Berkeley, Cal. (Feb. 10, 2009).

³⁸³ Interview with Franklin E. Zimring, Professor, Univ. of Cal., Berkeley Sch. of Law, in Berkeley, Cal. (Feb. 24, 2009).

³⁸⁴ *Id.*

³⁸⁵ *Id.*

³⁸⁶ *Id.*; See FRANKLIN E. ZIMRING ET AL., PUNISHMENT AND DEMOCRACY: THREE STRIKES AND YOU'RE OUT IN CALIFORNIA (2001).

³⁸⁷ Interview with Franklin E. Zimring, *supra* note 383.

³⁸⁸ *Id.*

³⁸⁹ *Id.*

it's administrative discretion, you win some and you lose some and there's no outside power."³⁹⁰

The vigor with which government agencies fight disclosures of offenders' DNA profiles suggests that researchers seeking to study them will lose an appeal to agency discretion.³⁹¹ Law enforcement agencies have an interest in secrecy³⁹²—courts analyzing DNA database searches under the Fourth Amendment have found them to be constitutional in part because DNA samples and profiles are not widely accessible.³⁹³ Furthermore, countless successful convictions rest on DNA statistics, which, if incorrect, could put those convictions at risk. At the very least, the discovery of incorrect statistics would bury courts in appeals and habeas corpus petitions. The state's desire to avoid this scenario is understandable, especially if DNA statistics turn out to be roughly but not precisely accurate. But avoidance is not acceptable. If law enforcement continues to champion adherence, without study, to current DNA statistics, control of the database should transfer to an independent agency.³⁹⁴ Given the incongruous reality that newborns' DNA samples are afforded fewer protections than offenders' profiles, the legislature should require DOJ to entrust this research to an outside state agency or scientist, or it should consider organizational changes that would prevent an agency with a vested interest in today's DNA statistics from being the database's gatekeeper.

In the meantime, judges should require government lawyers to demonstrate that law enforcement agencies have policies in place to allow independent statistical review of offender DNA databases. Courts should require government agencies to show that researchers in fact access offender profiles using those policies, and ensure that criminal defendants have access to the database as a source of exculpatory evidence. At the least, courts should allow defense counsel to present evidence questioning DNA rarity statistics

³⁹⁰ *Id.*

³⁹¹ *See supra* note 32 and surrounding text.

³⁹² *See* Giannelli, *supra* note 276.

³⁹³ *See* *United States v. Kriesel*, 508 F.3d 941, 944 (9th Cir. 2007) (pointing out the CODIS statute's "privacy protection standards"); *United States v. Weikert*, 504 F.3d 1, 13 (1st Cir. 2007) (noting the CODIS statute's criminal penalty for unauthorized use of offenders' DNA).

³⁹⁴ NAS makes a similar recommendation. *See* NAS REPORT, *supra* note 54, at 19.

routinely presented to juries. Without these steps, DNA evidence should be excluded from criminal trials.

CONCLUSION

Constitutional principles require government agencies to invite scientists to study offender DNA databases. State and federal law allow it. Just as public health concerns drive the limited release of newborns' genetic material, our system of justice and defendants' constitutional rights require the same for offenders' DNA profiles. Agencies cannot have it both ways: either the current system for releasing genetic information to researchers is adequate for forensic DNA profiles, or it is not adequate at all.

Ideally, statistical tests would show that current estimates as to the rarity of DNA profiles are accurate. However, if DNA profiles are not as rare as current statistics portray them to be, untold numbers of convictions could be compromised. This is not a reason to allow potentially faulty science to remain unexamined, it is a question that demands an answer. Yet no one outside government has access to the tools necessary to provide one. If law enforcement agencies continue to resist scrutiny of offender DNA profiles, reasonable judges should take action through the imperfect venue of a criminal trial. A better solution, however, would be to release offender DNA profiles to independent, qualified researchers applying appropriate safeguards.